ON STOCHASTIC SEA OF THE STANDARD MAP

ANTON GORODETSKI

ABSTRACT. Consider a generic one-parameter unfolding of a homoclinic tangency of an area preserving surface diffeomorphism. We show that for many parameters (residual subset in an open set approaching the critical value) the corresponding diffeomorphism has a transitive invariant set Ω of full Hausdorff dimension. The set Ω is a topological limit of hyperbolic sets and is accumulated by elliptic islands.

As an application we prove that stochastic sea of the standard map has full Hausdorff dimension for sufficiently large topologically generic parameters.

Keywords: standard map, conservative dynamics, hyperbolic set, Newhouse phenomena, persistent tangencies, Hausdorff dimension, homoclinic picture, homoclinic class, stochastic layer.

MSC 2000: 37E30, 37D45, 37J45, 37J20.

1. INTRODUCTION AND MAIN RESULTS

Here we prove that *stochastic sea of the Taylor-Chirikov standard map (i.e. the set of orbits with non-zero Lyapunov exponents) has full Hausdorff dimension for large topologically generic parameters.* In order to do so we show that *a perturbation of an area preserving diffeomorphism with a homoclinic tangency has hyperbolic invariant sets of almost full Hausdorff dimension.*

1.1. **Standard map.** The simplest and most famous symplectic system with highly non-trivial dynamics is the Taylor-Chirikov standard map of the two–dimensional torus $\mathbb{T}^2$, given by

$$(1) \qquad f_k(x,y) = (x + y + k\sin(2\pi x), y + k\sin(2\pi x)) \bmod \mathbb{Z}^2.$$

This family is related to numerous physical problems, see for example [C], [I], [SS]. For $k = 0$ the map $f_0(x,y) = (x+y,y)$ is completely integrable, and leaves the circles $\{y = \text{const}\}$ invariant. Due to KAM theory, after a perturbation large part of the torus is still foliated by invariant smooth circles, but lots of other phenomena appears: splitting of separatrices [GL], invariant Cantor sets [MMP], and many others [BL], [L], [Go]. Computer generated pictures show that chaotic part of the phase space (orbits with positive Lyapunov exponents) also form a subset of positive measure, but this was never rigorously justified. Due to Pesin's theory [P] this is equivalent to positivity of the metric entropy.

Main Question. (*Sinai* [Sin]) *Is the metric entropy of f_k positive for some values of k? for positive measure of values of k? for all non-zero values of k?*

A stronger version of this question is a famous conjecture which claims that the limit density at infinity of the set of parameters k for which the standard map $f_k : \mathbb{T}^2 \to \mathbb{T}^2$ is ergodic (and therefore has no elliptic islands) and non-uniformly hyperbolic with respect to Lebesgue measure is equal to

---

one. At the same time, it is known that the set of parameters k with this property (if non-empty) must be nowhere dense in a neighborhood of infinity [Du1].

In a more general way, one can ask (see [Bu], [X2]) whether an analytic symplectic map of a connected manifold can have coexisting chaotic component of positive measure and the Kolmogorov-Arnold-Moser (KAM) tori. There are C^∞ examples with this type of *mixed behavior* [Bu], [Do], [Li], [Pr], [W], but the rigorous proof of existence of mixed behavior for a real analytic map is still missing.

Our main result claims, roughly speaking, that stochastic sea of the standard map has full Hausdorff dimension for large topologically generic parameters.

Theorem 1. *There exists $k_0 > 0$ and a residual set $\mathcal{R} \in [k_0, +\infty)$ such that for every $k \in \mathcal{R}$ there exists an infinite sequence of transitive locally maximal hyperbolic sets of the map f_k*

$$(2) \qquad\qquad \Lambda_k^{(0)} \subseteq \Lambda_k^{(1)} \subseteq \Lambda_k^{(2)} \subseteq \ldots \subseteq \Lambda_k^{(n)} \subseteq \ldots$$

that has the following properties:

1. *The family of sets $\{\Lambda_k^{(0)}\}_{k \geq k_0}$ is dynamically increasing: for small $\varepsilon > 0$, $\Lambda_{k+\varepsilon}^{(0)}$ contains the continuation of $\Lambda_k^{(0)}$ at parameter $k + \varepsilon$;*

2. *The set $\Lambda_k^{(0)}$ is δ_k-dense in $\mathbb{T}^2$ for $\delta_k = \frac{4}{k^{1/3}}$;*

3. *Hausdorff dimension $\dim_H \Lambda_k^{(n)} \to 2$ as $n \to \infty$;*

4. *$\Omega_k = \overline{\cup_{n \in \mathbb{N}} \Lambda_k^{(n)}}$ is a transitive invariant set of the map f_k, and $\dim_H \Omega_k = 2$;*

5. *for any $x \in \Omega_k, k \in \mathcal{R}$, and any $\varepsilon > 0$ Hausdorff dimension*

$$\dim_H B_\varepsilon(x) \cap \Omega_k = \dim_H \Omega_k = 2,$$

where $B_\varepsilon(x)$ is an open ball of radius ε centered at x;

6. *Each point of Ω_k is an accumulation point of elliptic islands of the map f_k.*

The family of hyperbolic sets $\{\Lambda_k^{(0)}\}$ that satisfies properties 1. and 2. was constructed by Duarte in [Du1]. He also showed that $\dim_H \Lambda_k^{(0)} \to 2$ as $k \to \infty$, and that for topologically generic parameters the set $\Lambda_k^{(0)}$ is accumulated by elliptic islands.

For an open set of parameters our construction provides invariant hyperbolic sets of Hausdorff dimension arbitrarily close to 2.

Theorem 2. *There exists $k_0 > 0$ such that for any $\xi > 0$ there exists an open and dense subset $U \in [k_0, +\infty)$ such that for every $k \in U$ the map f_k has an invariant locally maximal hyperbolic set of Hausdorff dimension greater than $2 - \xi$ which is also δ_k-dense in $\mathbb{T}^2$ for $\delta_k = \frac{4}{k^{1/3}}$.*

Notice that these results give a partial explanation of the difficulties that we encounter studying the standard family. Indeed, one of the possible approaches is to consider an invariant hyperbolic set in the stochastic layer and to try to extend the hyperbolic behavior to a larger part of the phase space through homoclinic bifurcations. Unavoidably Newhouse domains (see [N1], [N2], [N4], [R] for dissipative case, and [Du2], [Du3], [Du4], [GS1] for the conservative case) associated with absence of hyperbolicity appear after small change of the parameter. If the Hausdorff dimension of the initial hyperbolic set is less than one, then the measure of the set of parameters that correspond to Newhouse

domains is small and has zero density at the critical value, see [NP], [PT2]. For the case when the Hausdorff dimension of the hyperbolic set is slightly bigger than one, similar result was recently obtained by Palis and Yoccoz [PY], and the proof is astonishingly involved. They also conjectured that analogous property holds for an initial hyperbolic set of any Hausdorff dimension, but the proof would require even more technical and complicated considerations.[1] Theorems 1 and 2 show that in order to understand the properties of the stochastic sea of the standard map using this approach one has to face these difficulties.

1.2. **Hyperbolic sets of large Hausdorff dimension.** Several famous long standing conjectures (including Main Question above) discuss the measure of certain invariant sets of some dynamical systems. Any set of positive Lebesgue measure has Hausdorff dimension which is equal to the dimension of the ambient manifold. Therefore it is reasonable to ask first whether those invariant sets indeed have full Hausdorff dimension.

In dissipative setting Downarowicz and Newhouse [DN] proved that there is a residual subset $\mathcal{R}$ of the space of C^r-diffeomorphisms of a compact two dimensional manifold M such that if $f \in \mathcal{R}$ and f has a homoclinic tangency, then f has compact invariant topologically transitive sets of Hausdorff dimension two. Their methods use essentially perturbative technics (see [GST]) and therefore cannot be generalized to the finite parameter families.

In conservative setting Newhouse [N5] proved that in $\mathrm{Diff}^1(M^2, Leb)$ there is a residual subset of maps such that every homoclinic class[2] for each of those maps has Hausdorff dimension 2. Later Arnaud, Bonatti and Crovisier [BC], [ABC] essentially improved that result and showed that in the space of C^1 symplectic maps the residual subset consists of the transitive maps that have only one homoclinic class (the whole manifold). Notice that due to KAM theory the low smoothness in that work is essential.

Here we show that *a generic one parameter area-preserving homoclinic bifurcation always give birth to a compact invariant topologically transitive set of Hausdorff dimension two.* This set is the closure of the union of a countable sequence of hyperbolic sets of Hausdorff dimension arbitrary close to two.

1.2.1. *The area preserving Henon family.* First of all we consider area preserving Henon family (3). For $a = -1$ this map has a degenerate fixed point at $(x, y) = (-1, 1)$. We construct invariant hyperbolic sets of large Hausdorff dimension for a slightly larger than -1 near this fixed point. Later we use the renormalization results to reduce the case of a generic unfolding of an area preserving surface diffeomorphism with a homoclinic tangency to this construction.

Theorem 3. *Consider the family of area preserving Henon maps*

$$(3) \qquad H_a : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y \\ -x + a - y^2 \end{pmatrix}.$$

---

[1]Here is a citation from Palis and Yoccoz [PY]: *"Of course, we expect the same to be true for all cases $0 < \dim_H(\Lambda) < 2$. For that, it seems to us that our methods need to be considerably sharpened: we have to study deeper the dynamical recurrence of points near tangencies of higher order (cubic, quartic, ...) between stable and unstable curves. We also hope that the ideas introduced in the present paper might be useful in broader contexts. In the horizon lies the famous question whether for the standard family of area preserving maps one can find sets of positive Lebesgue probability in parameter space such that the corresponding maps display non-zero Lyapunov exponents in sets of positive Lebesgue probability in phase space."*

[2]See Definition 1 below.

There is a (piecewise continuous) family of sets Λ_a, $a \in [-1, -1 + \varepsilon]$ for some $\varepsilon > 0$, such that the following properties hold.

   1. The set Λ_a is a locally maximal hyperbolic set of the map H_a;

   2. The set Λ_a contains a saddle fixed point of the map H_a;

   3. The set Λ_a has an open and closed (in Λ_a) subset $\widetilde{\Lambda}_a$ such that the first return map for $\widetilde{\Lambda}_a$ is a two-component Smale horseshoe;

   4. Hausdorff dimension $dim_H \widetilde{\Lambda}_a \to 2$ as $a \to -1$.

Theorem 3 should be considered as an improvement of Lemma A from [Du4], where Duarte proves that area preserving Henon maps have hyperbolic sets of large "left-right thickness" (see [Du4, Mo] for a definition) for values of a slightly larger than -1.

The proof essentially uses the construction from [Du2], [Du3] that was used by Duarte to study conservative Newhouse phenomena, and results regarding the splitting of separatrices for Henon family from [G1], [G2], [G3], [GSa], [BG] (see also [Ch], where some numerical results are described).

A similar statement holds also for any generic one parameter unfolding of an extremal periodic point (see [Du2] for a formal definition) as soon as the form of the splitting of separatrices can be established (see [G1, GL] for the relevant results on splitting of separatrices).

1.2.2. *Conservative homoclinic bifurcations and hyperbolic sets of large Hausdorff dimension.* In dissipative case Newhouse [N1] showed that near every surface diffeomorphism with a homoclinic tangency there are open sets (nowadays called *Newhouse domains*) of maps with persistence homoclinic tangencies. Moreover, in these open sets there are residual subsets of maps with infinitely many attracting periodic orbits. Later Robinson [R] showed that this result can be formulated in terms of generic one parameter unfolding of a homoclinic tangency.

In area preserving case Duarte [Du2], [Du3], [Du4] showed that homoclinic tangencies also lead to similar phenomena, the role of sinks is played by elliptic periodic points. Theorem 4 below is a stronger version of the Duarte's result: we can control the Hausdorff dimension of the hyperbolic sets that appear in the construction.

In order to construct transitive invariant sets of full Hausdorff dimension we use the notion of a *homoclinic class.*

Definition 1. *Let P be a hyperbolic saddle of a diffeomorphism f. A homoclinic class $H(P, f)$ is a closure of the union of all the transversal homoclinic points of P.*

It is known that $H(P, f)$ is a transitive invariant set of f, see [N3]. Moreover, consider all basic sets (locally maximal transitive hyperbolic sets) that contain the saddle P. A homoclinic class $H(P, f)$ is a smallest closed invariant set that contains all of them.

Theorem 4. *Let $f_0 \in \mathrm{Diff}^\infty(M^2, Leb)$[3] have an orbit $\mathcal{O}$ of quadratic homoclinic tangencies associated to some hyperbolic fixed point P_0, and $\{f_\mu\}$ be a generic unfolding of f_0 in $\mathrm{Diff}^\infty(M^2, Leb)$. Then for any $\delta > 0$ there is an open set $\mathcal{U} \subseteq \mathbb{R}^1$, $0 \in \overline{\mathcal{U}}$, such that the following holds:*

(1) for every $\mu \in \mathcal{U}$ the map f_μ has a basic set Δ_μ that contains the unique fixed point P_μ near P_0, exhibits persistent homoclinic tangencies, and Hausdorff dimension

$$\dim_H \Delta_\mu > 2 - \delta;$$

(2) there is a dense subset $\mathcal{D} \subseteq \mathcal{U}$ such that for every $\mu \in \mathcal{D}$ the map f_μ has a homoclinic tangency of the fixed point P_μ;

(3) there is a residual subset $\mathcal{R} \subseteq \mathcal{U}$ such that for every $\mu \in \mathcal{R}$

(3.1) the homoclinic class $H(P_\mu, f_\mu)$ is accumulated by f_μ's generic elliptic points,

(3.2) the homoclinic class $H(P_\mu, f_\mu)$ contains hyperbolic sets of Hausdorff dimension arbitrary close to 2; in particular, $\dim_H H(P_\mu, f_\mu) = 2$,

(3.3) $\dim_H \{x \in H(P_\mu, f_\mu) \mid P_\mu \in \omega(x) \cap \alpha(x)\} = 2$.

As usual, when we have a property that holds for a topologically generic parameter values, it is interesting to find out whether it holds for almost every parameter value, or with positive probability (i.e. for a positive measure set of parameters). For dissipative Newhouse phenomena see [TY], [GK2], [GHK] for some results in this direction. In the context of Theorem 4 this leads to the following questions.

Problem 1. *Under conditions of Theorem 4, what is the measure of the parameters μ such that $\dim_H H(P_\mu, f_\mu) = 2$? Such that $\dim_H H(P_\mu, f_\mu) > 2 - \varepsilon$?*

Also, for some applications (see [GK1]) it would be useful to improve the item (3.3) of Theorem 4.

Problem 2. *Under conditions of Theorem 4, prove that for every $\mu \in \mathcal{R}$ the set of points with dense orbits in the homoclinic class $H(P_\mu, f_\mu)$ has full Hausdorff dimension.*

Initially our interest in the conservative Newhouse phenomena was motivated by the fact that it appears in the three body problem. Namely, let us try to understand the structure of the set of oscillatory motions (a planet approaches infinity always returning to a bounded domain) in a Sitnikov problem [A, Si]. It is a special case of the restricted three body problem where the two primaries with equal masses are moving in an elliptic orbits of the two body problem, and the infinitesimal mass is moving on the straight line orthogonal to the plane of motion of the primaries which passes through the center of mass. The eccentricity of the orbits of primaries is a parameter. After some change of coordinates (McGehee transformation [McG]) the infinity can be considered as a degenerate saddle with smooth invariant manifolds that correspond to parabolic motions (the orbit tends to infinity with zero limit velocity). Stable and unstable manifolds coincide in the case of circular (parameter is equal to zero) Sitnikov problem. It is known that for non-zero eccentricity invariant manifolds have a point of transverse intersection [GP], [DH], [Mos]. This leads to the existence of homoclinic tangencies and appearance of all phenomena that can be encountered in the conservative homoclinic

---

[3]We assume C^∞-smoothness of diffeomorphisms here just for simplicity. For the renormalization procedures and arguments used in the current proof it is enough to assume only C^6-smoothness (which is probably not optimal either), compare with [Du4]. Since all the cases where we intend to apply this result (standard map, three body problems) are analytic, we are making no attempt to optimize the required class of smoothness.

bifurcations. Similar statement holds for the planar circular restricted three body problem. The existence of transversal homoclinic points in the latter case was established in [LS], [X1]. The farther development of this approach is a subject of our current joint project with V.Kaloshin, see [GK1] for some preliminary results.

The structure of the paper is the following. In Section 2 we remind the definitions of lateral (left- and right-) thickness of a Cantor set and show how Hausdorff dimension of a dynamically defined Cantor set can be estimated via its lateral thicknesses. In Section 3 Duarte's Distortion Theorem that allows to estimate thickness of a non-linear horseshoe is improved to cover a larger class of horseshoes. In Section 4 we discuss the results by Gelfreich and Sauzin [GSa] on splitting of separatrices in the area preserving Henon family, and then in Section 5 we apply those results together with results of Sections 2 and 3 to show how a horseshoe of large Hausdorff dimension appears (i.e. prove Theorem 3). In Section 6 we consider conservative homoclinic bifurcations and construct hyperbolic sets of large Hausdorff dimension (and prove Theorem 4), and, finally, in Section 7 we use this result to show that stochastic layer of the standard map has full Hausdorff dimension for many values of the parameter (i.e. prove Theorems 1 and 2).

2. Left-right thickness and Hausdorff dimension of Cantor sets

It is known that a Cantor set of large thickness must have large Hausdorff dimension [PT1]. In our construction we will encounter a Cantor set of small thickness. Nevertheless, we are still able to estimate Hausdorff dimension of the constructed Cantor sets. Namely, following the ideas by Moreira [Mo] and Duarte [Du3], we use lateral (left- and right-) thickness of a Cantor set, and we will see that the Cantor sets in our construction have one of the lateral thicknesses large and another one small (but controlled). In this section we show how to estimate Hausdorff dimension of a Cantor set in this case.

2.1. **Dynamically defined Cantor sets.** Here we reproduce the definition of the left and right thickness from [Du3] and [Mo]. We will use these one-sided thicknesses instead of the standard definition of thickness. See [PT1] for the usual definition of the thickness of a Cantor set.

Name *dynamically defined Cantor set* any pair (K, ψ) such that $K \subseteq \mathbb{R}$ is a Cantor set and $\psi : K \to K$ is a locally Lipschitz expanding map, topologically conjugated to some subshift of a finite type of a Bernoulli shift $\sigma : \{0, 1, \ldots, p\}^{\mathbb{N}} \to \{0, 1, \ldots, p\}^{\mathbb{N}}$. For the sake of simplicity, and because this is enough for our purpose, we will restrict ourselves to the case where ψ is conjugated to the full Bernoulli shift $\sigma : \{0, 1\}^{\mathbb{N}} \to \{0, 1\}^{\mathbb{N}}$. Also we will assume that a *Markov partition* $\mathcal{P} = \{K_0, K_1\}$ of (K, ψ) is given. In our case this means that the following properties are satisfied:

(1) $\mathcal{P}$ is a partition of K into disjoint union of two Cantor subsets, $K = K_0 \cup K_1$, $K_0 \cap K_1 = \emptyset$;

(2) the restriction of ψ to each K_i, $\psi|_{K_i} : K_i \to K$, is a strictly monotonous Lipschitz expanding homeomorphism.
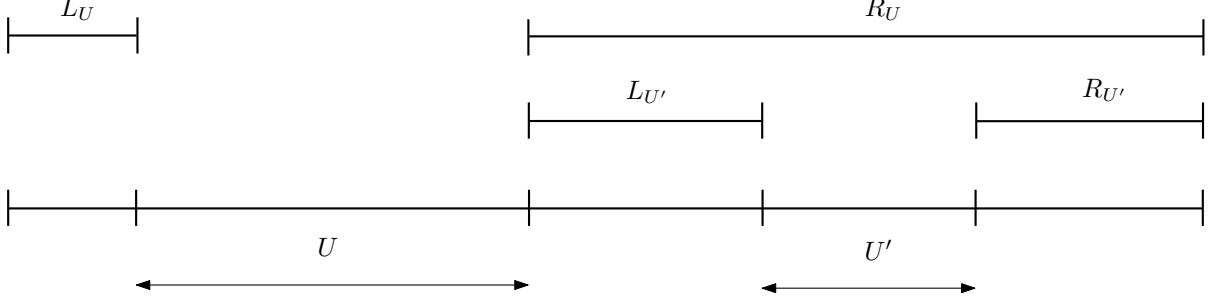
FIGURE 1. On definition of thickness

For a general definition of Markov partition see [Mo], [PT1].

Given a symbolic sequence $(a_0, \ldots, a_{n-1}) \in \{0,1\}^n$, denote

$$K(a_0, \ldots, a_{n-1}) = \cap_{i=0}^{n-1} \psi^{-i}(K_{a_i}),$$

then the map $\psi^n : K(a_0, \ldots, a_{n-1}) \to K$ is a Lipschitz expanding homeomorphism.

A bounded component of the complement $\mathbb{R} \backslash K$ is called a *gap* of K. For a dynamically defined Cantor set (K, ψ) the gaps are ordered in the following way. Denote by $\widehat{A}$ the convex hall of a subset $A \subseteq \mathbb{R}$. Then the interval $\widehat{K} \backslash (\widehat{K_0} \cup \widehat{K_1})$ is called a gap of order zero. A connected component of

$$\widehat{K} \backslash \cup_{(a_0, \ldots, a_{n-1}) \in \{0,1\}^n} \widehat{K(a_0, \ldots, a_{n-1})}$$

that is not a gap of order less than or equal to $n-1$ is called a gap of order n. It is straightforward to check that every gap of K is a gap of some finite order, and also that, given a gap $U = (x,y)$ of order n, for every $0 \le k \le n$ the open interval bounded by $\psi^k(x)$ and $\psi^k(y)$ is a gap of order $n-k$.

Definition 2. *Given a gap U of K with order n, we denote by L_U, respectively R_U, the unique interval of the form $\widehat{K(a_0, \ldots, a_{n-1})}$, with $(a_0, \ldots, a_{n-1}) \in \{0,1\}^n$, that is left, respectively right, adjacent to U. The greatest lower bounds*

$$\tau_L(K) = \inf \left\{ \frac{|L_U|}{|U|} : U \text{ is a gap of } K \right\}$$

$$\tau_R(K) = \inf \left\{ \frac{|R_U|}{|U|} : U \text{ is a gap of } K \right\}$$

are respectively called the left and right thickness of K. Similarly, the ratios

$$\tau_L(\mathcal{P}) = \frac{|L_{U_0}|}{|U_0|} \quad \text{and} \quad \tau_R(\mathcal{P}) = \frac{|R_{U_0}|}{|U_0|},$$

where U_0 is the unique gap of order zero, are called the left and the right thickness of the Markov partition $\mathcal{P}$.

Initially left- and right- thickness were introduced by Moreira[4] who proved the following generalization of the Newhouse's Gap Lemma [N1].

---

[4]Original Moreira's definition is formally different, and can be used for any Cantor set, not necessarily for dynamically defined. We use the modification suggested by Duarte in [Du3]. Lemma 1 holds in either case, and Duarte's definition is more convenient for dynamically defined Cantor sets.

Lemma 1. (Left-right gap lemma, see [Mo]) *Let (K^s, ψ^s), (K^u, ψ^u) be dynamically defined Cantor sets such that the intervals supporting K^s and K^u do intersect, K^s (resp. K^u) is not contained inside a gap of K^u (resp. K^s). If $\tau_L(K^s)\tau_R(K^u) > 1$ and $\tau_R(K^s)\tau_L(K^u) > 1$, then both Cantor sets intersect, $K^s \cap K^u \neq \emptyset$.*

2.2. **Large thickness implies large Hausdorff dimension.** Let us recall the definition of the Hausdorff dimension. Let $K \subset \mathbb{R}$ be a Cantor set and $\mathcal{U} = \{U_i\}_{i \in I}$ a finite covering of K by open intervals in $\mathbb{R}$. We define the diameter $\text{diam}(\mathcal{U})$ of $\mathcal{U}$ as the maximum of $|U_i|$, $i \in I$, where $|U_i|$ denotes the length of U_i. Define $H_\alpha(\mathcal{U}) = \sum_{i \in I} |U_i|^\alpha$. Then the *Hausdorff α-measure* of K is

$$m_\alpha(K) = \lim_{\varepsilon \to 0} \left(\inf_{\mathcal{U} \text{ covers } K, \text{ diam}(\mathcal{U}) < \varepsilon} H_\alpha(\mathcal{U}) \right).$$

It is not hard to see that there is a unique number, the *Hausdorff dimension* of K, denoted by $\dim_H(K)$, such that for $\alpha < \dim_H(K)$, $m_\alpha(K) = \infty$ and for $\alpha > \dim_H(K)$, $m_\alpha(K) = 0$.

Proposition 2. *Consider a Cantor set K, denote $\tau_L = \tau_L(K)$ and $\tau_R = \tau_R(K)$, and let d be the solution of the equation*

(4)
$$\tau_L^d + \tau_R^d = (1 + \tau_L + \tau_R)^d.$$

Then $\dim_H(K) \geq d$.

Remark 1. *One can consider Proposition 2 as a generalization of the Proposition 5 from Chapter 4.2 in [PT1], where the relation between the usual thickness and the Hausdorff dimension of a Cantor set was established. Indeed, if $\tau_L = \tau_R = \tau$ then Proposition 2 implies that $\dim_H(K) \geq \frac{\log 2}{\log\left(2 + \frac{1}{\tau}\right)}$, which is exactly the statement from [PT1].*

Proof of the Proposition 2. We will need the following elementary Lemma.

Lemma 3. *If $d \in (0, 1)$ is a solution of the equation (4) then*

$$\min \left\{ x^d + y^d \mid x \geq 0, y \geq 0, x + y \leq 1, x \geq \tau_L(1 - x - y), y \geq \tau_R(1 - x - y) \right\} = 1.$$

Proof of Lemma 3. The function $f(x, y) = x^d + y^d$ is concave, and takes value 1 at points $(0, 1), (1, 0)$, and $\left(\frac{\tau_L}{1 + \tau_L + \tau_R}, \frac{\tau_R}{1 + \tau_L + \tau_R} \right)$. Therefore its minimum in the triangle with vertices at these points is equal to 1. $\qquad\square$

We show that $H_d(\mathcal{U}) \geq (\text{diam } K)^d$ for every finite open covering $\mathcal{U}$ of K, which clearly implies the proposition. We can assume that $\mathcal{U}$ is a covering with disjoint intervals. This is no restriction because whenever two elements of $\mathcal{U}$ have nonempty intersection we can replace them by their union, getting in this way a new covering $\mathcal{V}$ such that $H_d(\mathcal{V}) \leq H_d(\mathcal{U})$. Note that, since $\mathcal{U}$ is an open covering of K, it covers all but finite number of gaps of K. Let U, a gap of K, have minimal order among the gaps of K which are not covered by $\mathcal{U}$. Let C^L and C^R be that bridges of K at the boundary points of U.

By construction there are A^L, $A^R \in \mathcal{U}$ such that $C^L \subset A^L$ and $C^R \subset A^R$. Take the convex hall A of $A^L \cup A^R$. Then

$$|A^L| \geq |C^L| \geq \tau_L \cdot |U| \geq \tau_L(|A| - |A^L| - |A^R|)$$

and

$$|A^R| \geq |C^R| \geq \tau_R \cdot |U| \geq \tau_R(|A| - |A^L| - |A^R|).$$

Or, equivalently,

$$\frac{|A^L|}{|A|} \geq \tau_L \left(1 - \frac{|A^L|}{|A|} - \frac{|A^R|}{|A|}\right)$$

and

$$\frac{|A^R|}{|A|} \geq \tau_R \left(1 - \frac{|A^L|}{|A|} - \frac{|A^R|}{|A|}\right).$$

Lemma 3 now implies that

$$\left(\frac{|A^L|}{|A|}\right)^d + \left(\frac{|A^R|}{|A|}\right)^d \geq 1,$$

and $|A^L|^d + |A^R|^d \geq |A|^d$. This means that the covering $\mathcal{U}_1$ of K obtained by replacing A^L and A^R by A in $\mathcal{U}$ is such that $H_d(\mathcal{U}_1) \leq H_d(\mathcal{U})$. Repeating the argument we eventually construct $\mathcal{U}_k$, a covering of the convex hall of K with $H_d(\mathcal{U}_k) \leq H_d(\mathcal{U})$. Since we must have $H_d(\mathcal{U}_k) \geq (\text{diam } K)^d$, this finishes the proof. $\qquad\square$

Proposition 2 can be used to find an explicit estimate of the Hausdorff dimension via one-sided thicknesses. In particular, when one of the one-sided thicknesses is very large and another one is small, the following Proposition gives an estimate that is good enough for our purposes.

Proposition 4. *Denote by τ_L and τ_R the left and right thicknesses of the Cantor set $K \subset \mathbb{R}$. Then*

$$dim_H K > \max\left(\frac{\log\left(1 + \frac{\tau_R}{1+\tau_L}\right)}{\log\left(1 + \frac{1+\tau_R}{\tau_L}\right)}, \frac{\log\left(1 + \frac{\tau_L}{1+\tau_R}\right)}{\log\left(1 + \frac{1+\tau_L}{\tau_R}\right)}\right).$$

Proof of Proposition 4. We will use the following Lemma.

Lemma 5. *Assume that for some $x, y > 0$, $x + y < 1$, and some $d_1, d_2 \in (0,1)$ the following relations hold:*

$$y = (1-x)^{\frac{1}{d_1}}, \qquad x^{d_2} + y^{d_2} = 1.$$

Then $d_2 > d_1$.

Proof of Lemma 5. Indeed, $(1-x)^{\frac{1}{d_1}} = y = (1 - x^{d_2})^{\frac{1}{d_2}} < (1-x)^{\frac{1}{d_2}}$, so due to our choice of x we have $d_2 > d_1$. $\qquad\square$
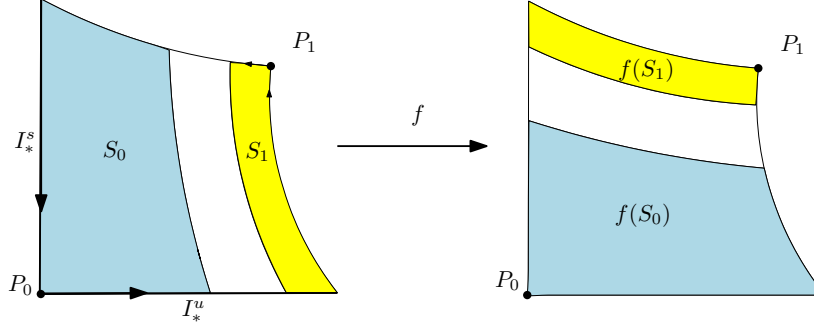
Let us apply Lemma 5 to $x = \frac{\tau_L}{1+\tau_L+\tau_R}$ and $y = \frac{\tau_R}{1+\tau_L+\tau_R}$. If $y = (1-x)^{\frac{1}{d_1}}, x^{d_2} + y^{d_2} = 1$ for some $d_1, d_2 \in (0,1)$ then by Proposition 2 we have

$$\dim_H K \geq d_2 > d_1 = \frac{\log(1-x)}{\log y} = \frac{\log\left(1 - \frac{\tau_L}{1+\tau_L+\tau_R}\right)}{\log\left(\frac{\tau_R}{1+\tau_L+\tau_R}\right)} = \frac{\log\left(1 + \frac{\tau_L}{1+\tau_R}\right)}{\log\left(1 + \frac{1+\tau_L}{\tau_R}\right)}.$$

In a similar way one can show that $\dim_H K > \frac{\log\left(1+\frac{\tau_R}{1+\tau_L}\right)}{\log\left(1+\frac{1+\tau_R}{\tau_L}\right)}.$ $\qquad\square$

Remark 2. *Assume that $\tau_R \sim \frac{1}{\lambda-1}$, $\tau_L \sim (\lambda-1)^\nu$. Then*

$$\lim_{\lambda \to 1+0} \frac{\log\left(1 + \frac{\tau_R}{1+\tau_L}\right)}{\log\left(1 + \frac{1+\tau_R}{\tau_L}\right)} = \lim_{\lambda \to 1+0} \frac{\log\left(1 + \frac{\frac{1}{\lambda-1}}{1+(\lambda-1)^\nu}\right)}{\log\left(1 + \frac{1+\frac{1}{\lambda-1}}{(\lambda-1)^\nu}\right)} = \frac{1}{1+\nu}.$$

FIGURE 2. A nonlinear horseshoe $f \in \mathfrak{F}$

So if ν is small enough and λ is close to one, then $\dim_H K$ is close to 1.

3. Nonlinear horseshoes and distortion estimates

Here we describe the way to estimate lateral thicknesses of a non-linear horseshoe. We follow the approach from [Du3], but with some modifications; since our goal is to construct horseshoes with large (close to 2) Hausdorff dimension, we have to deal with a larger class of horseshoes than the one considered in [Du3].

3.1. Non-linear horseshoes and their Markov partitions.

Definition 3. *Define $\mathcal{F}$ to be the set of all maps $f : \mathbf{S}_0 \cup \mathbf{S}_1 \to \mathbb{R}^2$ such that:*

(1) $\mathbf{S}_0$, $\mathbf{S}_1 \subset \mathbb{R}^2$ are compact sets, diffeomorphic to rectangles, with non-empty interior;

(2) f is a map of class C^2, in a neighborhood of $\mathbf{S}_0 \cup \mathbf{S}_1$, mapping this compact set diffeomorphically onto its image $f(\mathbf{S}_0) \cup f(\mathbf{S}_1)$;

(3) the maximal invariant set $\Lambda(f) = \cap_{n \in \mathbb{Z}} f^{-n}(\mathbf{S}_0 \cup \mathbf{S}_1)$ is a hyperbolic basic set conjugated to the topological Bernoulli shift $\sigma : \{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{Z}}$;

(4) $\mathcal{P} = \{\mathbf{S}_0, \mathbf{S}_1\}$ is a Markov partition for $f : \Lambda(f) \to \Lambda(f)$, in particular, f has two fixed points, $\mathbf{P}_0 \in \mathbf{S}_0$ and $\mathbf{P}_1 \in \mathbf{S}_1$, whose stable and unstable manifolds contain the boundaries of $\mathbf{S}_0$ and $\mathbf{S}_1$;

(5) both fixed points $\mathbf{P}_0$ and $\mathbf{P}_1$ have positive eigenvalues.

The action of f and f^{-1} respectively on the stable, and unstable, foliation of Λ,

$$\mathcal{F}^s = \{\text{connected comp. of } W^s(\Lambda) \cap (\mathbf{S}_0 \cup \mathbf{S}_1)\},$$

$$\mathcal{F}^u = \{\text{connected comp. of } W^u(\Lambda) \cap (f(\mathbf{S}_0) \cup f(\mathbf{S}_1))\},$$

can be described in the following way. Define

$$I_*^s = W_{loc}^s(\mathbf{P}_0) \cap \mathbf{S}_0 \quad \text{and} \quad I_*^u = W_{loc}^u(\mathbf{P}_0) \cap f(\mathbf{S}_0).$$

I_*^s and I_*^u are stable and unstable leaves of Λ respectively transversal to the foliation $\mathcal{F}^u$ and $\mathcal{F}^s$. Then the Cantor sets

$$K^s = \Lambda \cap I_*^u \quad \text{and} \quad K^u = \Lambda \cap I_*^s,$$

can be identified with the set of stable leaves of $\mathcal{F}^s$, respectively unstable leaves of $\mathcal{F}^u$. Define the projections $\pi_s : \Lambda \to K^s$ and $\pi_u : \Lambda \to K^u$ in the obvious way: $\pi_s(P)$ is the unique point in $W^s_{loc}(P) \cap I^u_*$, and similarly $\pi_u(P)$ is the unique point in $W^u_{loc}(P) \cap I^s_*$. The maps $\psi^s : K^s \to K^s$ and $\psi^u : K^u \to K^u$,

$$\psi^s = \pi_s \circ f \qquad \text{and} \qquad \psi^u = \pi_u \circ f^{-1},$$

describe the action of f, respectively f^{-1}, on stable, respectively unstable leaves of Λ. The pairs (K^s, ψ^s) and (K^u, ψ^u) are dynamically defined Cantor sets, topologically conjugated to the Bernoulli shift $\sigma : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$, with Markov partitions $\mathcal{P}^u = \{I^u_* \cap \mathbf{S}_0, I^u_* \cap \mathbf{S}_1\}$ and $\mathcal{P}^s = \{I^s_* \cap f(\mathbf{S}_0), I^s_* \cap f(\mathbf{S}_1)\}$.

3.2. Distortion of a dynamically defined Cantor set.

Definition 4. *Given a Lipschitz expanding map $g : J \to \mathbb{R}$, defined on some subset $J \subset \mathbb{R}$, we define distortion of g on J in the following way:*

$$\text{Dist}(g, J) = \sup_{x,y,z \in J} \log \left\{ \frac{|g(y) - g(x)|}{|g(z) - g(x)|} \frac{|z - x|}{|y - x|} \right\} \in [0, +\infty],$$

where the sup is taken over all $x, y, z \in J$ such that $z \neq x$ and $y \neq x$; due to injectivity of g this implies that $g(z) \neq g(x)$ and $g(y) \neq g(x)$.

Reversing the roles of y and z we see that the distortion is always greater than or equal to $\log 1 = 0$. If $\text{Dist}(g, J) = c$, then for all $x, y, z \in J$ with $z \neq x$ and $y \neq x$ we have

$$e^{-c} \frac{|y - x|}{|z - x|} \leq \frac{|g(y) - g(x)|}{|g(z) - g(x)|} \leq e^c \frac{|y - x|}{|z - x|}.$$

Definition 5. *The distortion of a dynamically defined Cantor set (K, ψ) is defined as*

$$\text{Dist}_\psi(K) = \sup \text{Dist}(\psi^n, K(a_0, \ldots, a_{n-1}))$$

taken over all sequences $(a_0, \ldots, a_{n-1}) \in \{0,1\}^n$.

Lemma 6 (see [PT1], [Du3]). *Let (K, ψ) be a dynamically defined Cantor set with a Markov partition $\mathcal{P}$ and distortion $\text{Dist}_\psi(K) = c$. Then*

$$e^{-c} \tau_L(\mathcal{P}) \leq \tau_L(K) \leq e^c \tau_L(\mathcal{P}), \qquad e^{-c} \tau_R(\mathcal{P}) \leq \tau_R(K) \leq e^c \tau_R(\mathcal{P}).$$

3.3. Duarte's Distortion Theorem.
For $C^* = 2$ the following Definition coincides with Definition 4 from [Du3].

Definition 6. *Given positive constants C^* along with small ε and γ, define $\mathcal{F}(C^*, \varepsilon, \gamma)$ to be the class of all maps $f : \mathbf{S}_0 \cup \mathbf{S}_1 \to \mathbb{R}^2$, $f \in \mathcal{F}$, such that:*

(1) $\text{diam}(\mathbf{S}_0 \cup \mathbf{S}_1) \leq 1, \text{diam}(f(\mathbf{S}_0) \cup f(\mathbf{S}_1)) \leq 1$;

(2) *the derivative of f, $Df_{(x,y)} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where a, b, c and d are C^1-functions, satisfies all over*
$\mathbf{S}_0 \cup \mathbf{S}_1$

(a) $\det Df = ad - bc = 1$,

(b) $|d| < 1 < |a| \leq C^*/\varepsilon$,

(c) $|b|, |c| \leq \varepsilon(|a| - 1)$;

(3) the C^1-functions on $f(\mathbf{S}_0) \cup f(\mathbf{S}_1)$, $\tilde{a} = a \circ f^{-1}$, $\tilde{b} = b \circ f^{-1}$, $\tilde{c} = c \circ f^{-1}$ and $\tilde{d} = d \circ f^{-1}$, i.e.
$$Df^{-1}_{(x,y)} = \begin{pmatrix} \tilde{d} & -\tilde{b} \\ -\tilde{c} & \tilde{a} \end{pmatrix}, \text{ satisfy}$$

(a) $\left| \frac{\partial \tilde{b}}{\partial x} \right| = \left| \frac{\partial \tilde{d}}{\partial y} \right|, \left| \frac{\partial \tilde{b}}{\partial y} \right|, \left| \frac{\partial \tilde{c}}{\partial x} \right|, \left| \frac{\partial \tilde{a}}{\partial x} \right| = \left| \frac{\partial \tilde{c}}{\partial y} \right| \leq \gamma(|\tilde{a}| - 1),$

(b) $\left| \frac{\partial a}{\partial y} \right| = \left| \frac{\partial b}{\partial x} \right|, \left| \frac{\partial b}{\partial y} \right|, \left| \frac{\partial c}{\partial x} \right|, \left| \frac{\partial c}{\partial y} \right| = \left| \frac{\partial d}{\partial x} \right| \leq \gamma(|a| - 1),$

(c) $\left| \frac{\partial \tilde{a}}{\partial y} \right|, \left| \frac{\partial \tilde{d}}{\partial x} \right| \leq \gamma|\tilde{a}|(|\tilde{a}| - 1),$

(d) $\left| \frac{\partial a}{\partial x} \right|, \left| \frac{\partial d}{\partial y} \right| \leq \gamma|a|(|a| - 1);$

(4) the variation of $\log |a(x,y)|$ in each rectangle S_i is less or equal to $\gamma(1 - \alpha_i^{-1})$, where $\alpha_i = \max_{(x,y) \in S_i} |a(x,y)|$;

(5) finally, the gap sizes satisfy:
$$\text{dist}\,(\mathbf{S}_0 \cup \mathbf{S}_1) \geq \frac{\varepsilon}{\gamma} \qquad and \qquad \text{dist}\,(f(\mathbf{S}_0), f(\mathbf{S}_1)) \geq \frac{\varepsilon}{\gamma}.$$

The nice feature of the maps from $\mathcal{F}(C^*, \varepsilon, \gamma)$ is that the stable and unstable foliations have small uniformly bounded distortion, as the following theorem shows.

Theorem 5. *For a given $C^* > 0$ and all small enough $\varepsilon > 0$ and $\gamma > 0$, given $f \in \mathcal{F}(C^*, \varepsilon, \gamma)$, the basic set $\Lambda(f)$ gives dynamically defined Cantor sets (K^u, ψ^u) and (K^s, ψ^s) with distortion, bounded by $D(C^*, \varepsilon, \gamma) = 4(C^* + 3)\gamma + 2\varepsilon$. In particular,*

$$e^{-D(C^*,\varepsilon,\gamma)} \tau_L(\mathcal{P}^s) \leq \tau_L(K^s(f)) \leq e^{D(C^*,\varepsilon,\gamma)} \tau_L(\mathcal{P}^s),$$

$$e^{-D(C^*,\varepsilon,\gamma)} \tau_R(\mathcal{P}^s) \leq \tau_R(K^s(f)) \leq e^{D(C^*,\varepsilon,\gamma)} \tau_R(\mathcal{P}^s),$$

$$e^{-D(C^*,\varepsilon,\gamma)} \tau_L(\mathcal{P}^u) \leq \tau_L(K^u(f)) \leq e^{D(C^*,\varepsilon,\gamma)} \tau_L(\mathcal{P}^u),$$

$$e^{-D(C^*,\varepsilon,\gamma)} \tau_R(\mathcal{P}^u) \leq \tau_R(K^u(f)) \leq e^{D(C^*,\varepsilon,\gamma)} \tau_R(\mathcal{P}^u).$$

Remark 3. *Again, for $C^* = 2$ this Theorem coincides with Theorem 2 in [Du3]. Notice that conditions (2b) and (2c) of definition 6 imply that $|b|, |c| \leq C^*$, and for $C^* = 2$ this gives an unreasonable restriction on the class of maps that could be considered. We will need to apply Theorem 5 for a map which belongs to the class $\mathcal{F}(C^*, \varepsilon, \gamma)$ with larger value of C^*, see Proposition 20.*

Proof of Theorem 5. The straightforward repetition of the proof of Theorem 2 in [Du3] with the necessary adjustments needed to take the constant C^* into account proves Theorem 5. The only place in the proof of Theorem 2 from [Du3] where the condition $|a| \leq \frac{2}{\varepsilon}$ is used is the inequality (3) from Lemma 4.2. If we use the inequality $|a| \leq \frac{C^*}{\varepsilon}$ instead, 6γ should be replaced by $\frac{3}{2}(C^* + 2)$ there. Due to this change, in Lemma 4.1 from [Du3] one should take $2(C^* + 2)\gamma$ instead of 8γ as an upper bound of Lipschitz seminorm $Lip(\sigma^s)$ and $Lip(\sigma^u)$ of functions σ^s and σ^u that describe stable and unstable foliations. This leads to similar changes in the statement of Lemma 4.4 from [Du3] and in the estimate of the distortion. Finally we use Lemma 6 to finish the proof of Theorem 5. $\square$

4. Birkhoff and Gelfreich normal forms

In this section we collect some known results on quadratic families, Birkhoff normal form, and splitting of separatrices, in many cases reformulating them to adapt to our setting, so preparing to use them later in Section 5.

4.1. Uniqueness of the area preserving quadratic family.

The non-trivial quadratic diffeomorphism of the plane which preserve area and orientation with a fixed point can be put after a linear change of coordinates into the Henon family (3), see [H]. In particular, we can consider the family

$$(5) \qquad F_\varepsilon : (x, y) \mapsto (x + y - x^2 + \varepsilon,\ y - x^2 + \varepsilon)$$

instead of (3). In this form it is a partial case of a so called *generalized standard family*, and it was considered in [G1].

Another form of the quadratic area preserving family[5], namely

$$(6) \qquad \widetilde{F}_\varepsilon : (x, y) \mapsto (x + \varepsilon(y + \varepsilon(x - x^2)),\ y + \varepsilon(x - x^2)),$$

was used in [GSa].

4.2. Rescaling and the family of maps close to identity.

Let us start with the family F_ε (5). Consider the following family of the affine coordinate changes:

$$\Upsilon_\delta \begin{pmatrix} u \\ v \end{pmatrix} = - \begin{pmatrix} \delta^2 \\ 0 \end{pmatrix} + \begin{pmatrix} \delta^2 & 0 \\ 0 & \delta^3 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix},$$

where $\delta = \varepsilon^{\frac{1}{4}}$. Then

$$\Upsilon_\delta^{-1} \circ F_{\delta^4} \circ \Upsilon_\delta \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u + \delta v \\ v + \delta(2u - u^2) \end{pmatrix} + \delta^2 \begin{pmatrix} 2u - u^2 \\ 0 \end{pmatrix}.$$

Now we have a family of area preserving maps close to identity. For each of these maps the origin is a saddle with eigenvalues

$$\lambda_1 = 1 + \delta^2 + \sqrt{\delta^4 + 2\delta^2} = 1 + \sqrt{2}\delta + O(\delta^2) > 1,$$

$$\lambda_2 = \lambda_1^{-1} = 1 + \delta^2 - \sqrt{\delta^4 + 2\delta^2} = 1 - \sqrt{2}\delta + O(\delta^2) < 1.$$

Set $h = \log \lambda_1$. By definition $h = \sqrt{2}\delta + O(\delta^2)$, and δ can be given by implicit function of h. Define the following (rescaled and reparametrized) family

$$(7) \qquad \mathfrak{F}_h : (u, v) \mapsto (u, v) + \delta(v, 2u - u^2) + \delta^2(2u - u^2, 0),\ \delta = \delta(h).$$

---

[5]See Section 4 in [Du4] for an explicit change of coordinates sending the family $\widetilde{F}_\varepsilon$ to the Henon family (3). Notice also that in Subsection 4.2 the family (5) is send to the family (7) by an affine change of coordinates, and the change of coordinates $(u, v) \mapsto (2x, 2\sqrt{2}y)$ together with reparametrization $\varepsilon = \sqrt{2}\delta$ conjugates the family (7) with the family (6).

4.3. **Birkhoff normal form.** A real analytic area preserving diffeomorphism of a two dimensional domain in a neighborhood of a saddle with eigenvalues (λ, λ^{-1}) by an analytic change of coordinate can be reduced to the Birkhoff normal form ([S], see also [SM]):

$$(8) \qquad\qquad\qquad N(x,y) = (\Delta(xy)x, \Delta^{-1}(xy)y),$$

where $\Delta(xy) = \lambda + a_1 xy + a_2(xy)^2 + \dots$ is analytic.

We need a generalization of this Birkhoff normal form for one-parameter families. In the following statement we combine the results of Proposition 3.1 from [FS1] and of Proposition 4 from [Du4].

Theorem 6. *There exists a neighborhood U of the origin such that for all $h \in (0, h_0)$ there exists a coordinate change $C_h : U \to \widehat{U}$ with the following properties:*

1. If $N_h = C_h \mathfrak{F}_h C_h^{-1}$ then $N_h(u, v) = (\Delta_h(uv)u, \Delta_h^{-1}(uv)v)$, where $\Delta_h(uv) = \lambda(h) + a_1(h)uv + a_2(h)(uv)^2 + \dots$ is analytic.

2. C^3-norms of the coordinate changes C_h are uniformly bounded with respect to the parameter h.

3. $\Delta_h(s) \geq 1$ is a smooth function of s and h.

Remark 4. *The second property is not formulated explicitly in [FS1] or in [Du4] but it immediately follows from Cauchy estimates. Indeed, it follows from the proof there that the map C_h is analytic and radius of convergence of the corresponding series is uniformly bounded from below.*

Also we will need the following property of the parametric Birkhoff normal form for the family $\mathfrak{F}_h$.

Lemma 7. *For some constant $C > 0$ and small enough $h_0 > 0$ and $s_0 > 0$ the following holds. For all $h \in [0, h_0)$ and $s \in [0, s_0)$*

1. $\log \Delta_h(s) \geq C^{-1}h$,

2. $|\Delta_h'(s)| \leq Ch$,

*3. $|\Delta_h''(s)| \leq Ch$.*

Remark 5. *This Lemma is similar to Lemma 6.3 from [Du3], but in our case we have one, not two parameter family, and therefore those two statements are essentially different.*

Proof of Lemma 7. Consider $g(s, h) = \log \Delta_h(s)$. We have $g(s, 0) = 0$, $g(0, h) = h$, and g is a smooth function of (s, h). This implies that for small enough $s_0 > 0$, $h_0 > 0$ and large $C > 0$ we have $g(s, h) \geq C^{-1}h$ for all $s \in [0, s_0]$ and $h \in [0, h_0]$.

From the explicit form of the family $\mathfrak{F}_h$ (7) we see that $\mathfrak{F}_h \to \mathrm{Id}$ as $h \to 0$ in C^r-norm for every $r \in \mathbb{N}$. Since C^3-norms of C_h and C_h^{-1} are uniformly bounded, this implies that $N_h \to \mathrm{Id}$ in C^2-norm as $h \to 0$. In particular,

$$DN_h(x,y) = \begin{pmatrix} \Delta_h(xy) + \Delta_h'(xy)xy & x^2\Delta_h'(xy) \\ -\frac{\Delta_h'(xy)y^2}{\Delta_h^2(xy)} & \Delta_h^{-1}(xy) - \frac{\Delta_h'(xy)xy}{\Delta^2(xy)} \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

as $h \to 0$ uniformly in $(x, y) \in U$. This implies that $\Delta_h'(s) \to 0$ as $h \to 0$ uniformly in $s \in [s_1, s_0]$ for every $s_1 \in (0, s_0)$. Also $\Delta_0(s) = 1$ for every $h \in [0, h_0)$, so $\Delta_0'(s) = 0$. Since $\Delta_h'(s)$ is a continuous function, this implies that $\Delta_h'(s) \to 0$ as $h \to 0$ uniformly in $s \in [0, s_0]$. Since $\Delta_h'(s)$ is a smooth function of (s, h), this implies that $|\Delta_h'(s)| \leq Ch$ if $C > 0$ is large enough. Similarly one can show that $|\Delta_h''(s)| \leq Ch$. $\qquad\square$
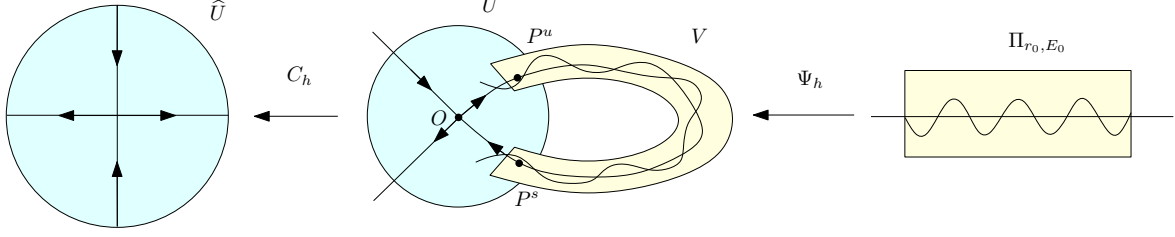
FIGURE 3. Birkhoff and Gelfreich normal forms for $\mathfrak{F}_h$

4.4. **Gelfreich normal form and splitting of separatrices.** The family $\mathfrak{F}_h$ is closely related to the conservative vector field

$$(9) \qquad \begin{cases} \dot{x} = y, \\ \dot{y} = 2x - x^2. \end{cases}$$

Namely, due to Theorems A and A′ from [FS2] (see also Proposition 5.1 from [FS1]) the separatrix phase curve of the vector field (9) (let us denote it by σ) gives a good approximation of some finite pieces of $W^s_{\mathfrak{F}_h}(0,0)$ and $W^u_{\mathfrak{F}_h}(0,0)$. Denote by $\widetilde{\sigma}$ a segment of separatrix σ that contains some points $P^u \in W^u_{loc}(0,0) \cap U$ and $P^s \in W^s_{loc}(0,0) \cap U$ and by V a neighborhood of $\widetilde{\sigma}$. Denote by $\widetilde{W}^s_h(0,0)$ the finite piece of $W^s_{\mathfrak{F}_h}(0,0)$ between the points where $W^s_{\mathfrak{F}_h}(0,0)$ leaves U for the first time and the first point where $W^s_{\mathfrak{F}_h}(0,0)$ returns to U again. Define $\widetilde{W}^u_h(0,0)$ in a similar way. Then $\widetilde{W}^s_h(0,0)$ and $\widetilde{W}^u_h(0,0)$ are always (for all $h \in (0, h_0)$) in V.

The restriction of the map $\mathfrak{F}_h$ on the local separatrix $W^u_{loc}(0,0)$ is conjugated with a multiplication $\xi \mapsto \lambda\xi$, $\xi \in (\mathbb{R}, 0)$. Let us call a parameter t on $W^u_{\mathfrak{F}_h}(0,0)$ *standard* if it is obtained by a substitution of e^t instead of ξ into the conjugating function. Such a parametrization is defined up to a substitution $t \mapsto t + const$.

Denote $\Pi_{r_0, E_0} = \{(t, E) \in \mathbb{R}^2 \mid |t| < r_0, |E| < E_0\}$.

In the following Theorem we summarize the results from [G1, G2, G3, GSa, BG].

Theorem 7. *There are neighborhood V of the segment of σ between points P^u and P^s and constants r_0 and E_0 such that for some $h_0 > 0$ and all $h \in (0, h_0)$ there exists a map $\Psi_h : \Pi_{r_0, E_0} \to \mathbb{R}^2$ with the following properties:*

1. $\Psi_h(\Pi_{r_0, E_0}) \supset V$;

2. Ψ_h *is real analytic;*

3. Ψ_h *is area preserving;*

4. Ψ_h *conjugates the map $\mathfrak{F}_h$ with the shift $H_h : (t, E) \mapsto (t + h, E)$;*

5. *C^3-norms of Ψ_h and Ψ_h^{-1} are uniformly bounded with respect to $h \in (0, h_0)$.*

6. $\Psi_h^{-1}(\widetilde{W}^u_h) = \{E = 0\}$, *and t gives a standard parametrization of the unstable manifold;*

7. Stable manifold $\Psi_h^{-1}(\widetilde{W}_h^s)$ can be represented as a graph of a real-analytic h-periodic function $\Theta(t)$ such that

$$\Theta(t) = 8\sqrt{2}|\Theta_1| h^{-6} e^{-2\pi^2/h} \sin\frac{2\pi t}{h} + O(h^{-5} e^{-2\pi^2/h}); \tag{10}$$

$$\dot\Theta(t) = 16\sqrt{2}\pi|\Theta_1| h^{-7} e^{-2\pi^2/h} \cos\frac{2\pi t}{h} + O(h^{-6} e^{-2\pi^2/h}); \tag{11}$$

$$\ddot\Theta(t) = -32\sqrt{2}\pi^2|\Theta_1| h^{-8} e^{-2\pi^2/h} \sin\frac{2\pi t}{h} + O(h^{-7} e^{-2\pi^2/h}); \tag{12}$$

8. $|\Theta_1| \neq 0$.

In [G3] existence of the normal form Ψ_h that satisfies properties 1.-6. was shown. In [G1, G2] the splitting of separatrices (the form of the function $\Theta(t)$) was studied for the initial family F_ε (5). The fact that $|\Theta_1| \neq 0$ was proved in [GSa]. In the recent paper [BG] the whole asymptotic series for $\Theta(t)$ is presented (in fact, for a much wider class of families that includes area preserving Henon family), but here we stated only the partial case of that result which we will need in Section 5.

Remark 6. *To simplify the notation define the function*

$$\mu(h) = 16\sqrt{2}\pi|\Theta_1| h^{-7} \exp\left(-2\pi^2/h\right). \tag{13}$$

Notice that the angle between $\widetilde{W}_h^u$ and $\widetilde{W}_h^s$ at the homoclinic point in the normalized coordinates is equal to $\mu(h)(1 + O(h))$. The function $\Theta(t)$ can now be represented in the following way:

$$\Theta(t) = \frac{1}{2\pi}h\mu(h)\sin\frac{2\pi t}{h} + O(h^2\mu(h)), \quad \dot\Theta(t) = \mu(h)\cos\frac{2\pi t}{h} + O(h\mu(h)),$$

$$\ddot\Theta(t) = -2\pi h^{-1}\mu(h)\sin\frac{2\pi t}{h} + O(\mu(h)).$$

5. Construction of a thick horseshoe for area preserving Henon family

Here we prove Theorem 3. In order to do so we construct the horseshoe for the first return map in a neighborhood of a saddle, verify the conditions of the Duarte's Distortion Theorem from Section 3, and obtain estimates on one-sided thicknesses of the constructed horseshoe. Relations between one-sided thicknesses and Hausdorff dimension obtained in Section 2 will imply the required result.

5.1. **Construction of the domain for the first return map.** Let q_h^u be the closest to $P^u \in \sigma$ point of intersection of $\widetilde{W}_h^u$ and $\widetilde{W}_h^s$. Consider a finite sequence of images of q_h^u under the map $\mathfrak{F}_h$ that belong to the neighborhood V, $\{q_h^u, \mathfrak{F}(q_h^u), \mathfrak{F}^2(q_h^u), \ldots\}$. Let q_h^s be the point of this sequence closest to the point $P^s \in \sigma$. Define $k(h) \in \mathbb{N}$ by $\mathfrak{F}_h^{k(h)}(q_h^u) = q_h^s$. Take the vector $v = (1, 0) \in T_{C_h(q_h^u)}\widehat{U}$ and consider the vector $w = (w_1, w_2) = D(C_h \circ \Psi_h \circ H^{k(h)} \circ \Psi_h^{-1} \circ C_h^{-1})v \in T_{C_h(q_h^s)}\widehat{U}$. Without loss of generality we can assume that $w_1 > 0$ (otherwise just take a homoclinic point between q_h^u and $\mathfrak{F}(q_h^u)$ instead of q_h^u). Scaling, if necessary, we can assume that in the Birkhoff normalizing coordinates we have $C_h(q_h^u) = (1, 0)$, $C_h(q_h^s) = (0, 1)$.

Fix small $\nu > 0$. Recall that $\lambda = \Delta_h(0) = e^h$. Set

$$n = \left[-\frac{\log(\mu(h)h^{1+\nu})}{2h}\right]. \tag{14}$$

Due to this choice $\lambda^{-2n} \approx \mu(h)h^{1+\nu}$. More precisely, $\lambda^{-2n} \in [\mu(h)h^{1+\nu}, \lambda^2\mu(h)h^{1+\nu})$.
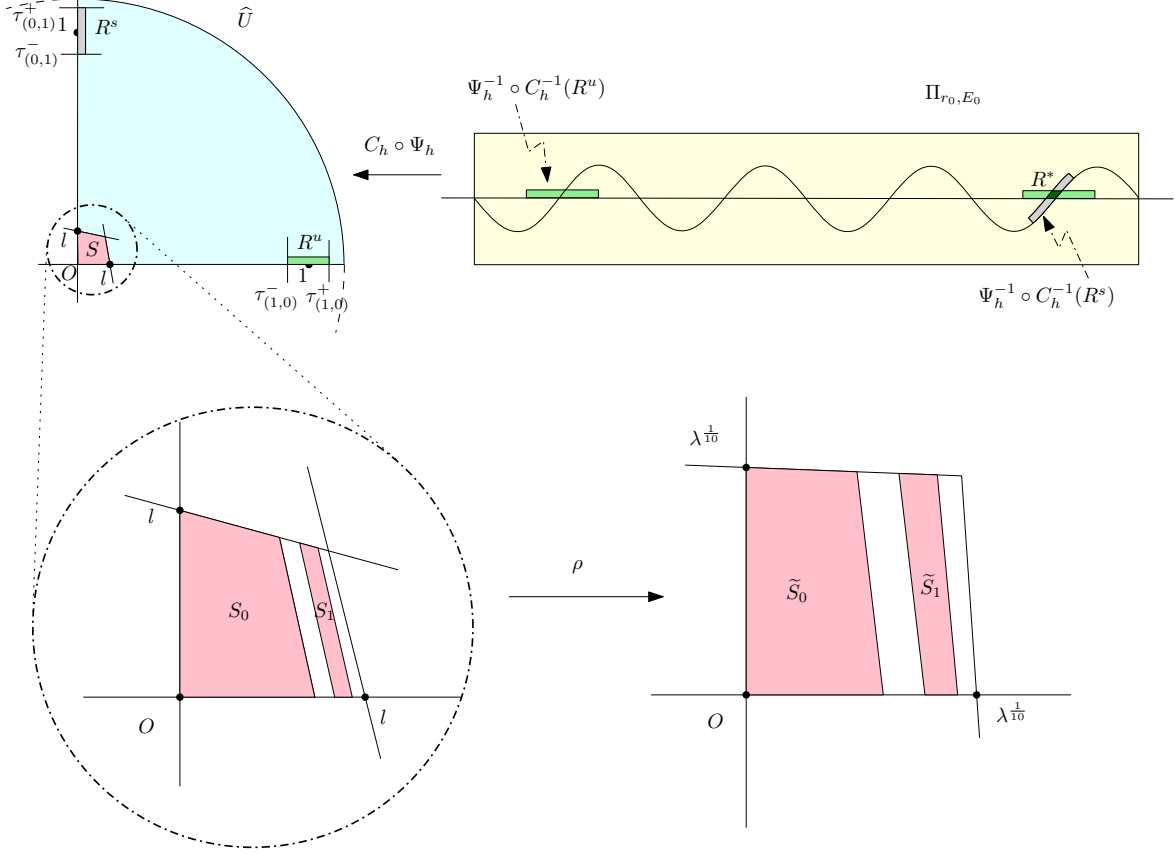
FIGURE 4. Construction of the horseshoe.

Remark 7. *Notice that this choice of n for $\nu = \frac{1}{2}$ is analogous to the formula (7) in* [Du3].

Define the following lines:

$$\tau_{(1,0)}^+ = \{x = \lambda^{\frac{1}{10}}\}, \quad \tau_{(1,0)}^- = \{x = \lambda^{-\frac{1}{10}}\},$$
$$\tau_{(0,1)}^+ = \{y = \lambda^{\frac{1}{10}}\}, \quad \tau_{(0,1)}^- = \{y = \lambda^{-\frac{1}{10}}\}.$$

Denote by S the square formed by coordinate axes, $N_h^n(\tau_{(0,1)}^+)$, and $N_h^{-n}(\tau_{(1,0)}^+)$. The bottom and left edges of S have the size

(15) $$l = \lambda^{-n+\frac{1}{10}}.$$

Notice that since DN_h is close to the linear map $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$, the curve $N_h^n(\tau_{(0,1)}^+)$ (resp., $N_h^{-n}(\tau_{(1,0)}^+)$) is C^1-close to a horizontal (resp., vertical) line. Denote by R^u and R^s the rectangles formed by x-axis, $\tau_{(1,0)}^+$, $\tau_{(1,0)}^-$ and $N_h^{2n}(\tau_{(0,1)}^+)$, and by y-axis, $\tau_{(0,1)}^+$, $\tau_{(0,1)}^-$ and $N_h^{-2n}(\tau_{(1,0)}^+)$, respectively. Notice that $R^u = N_h^n(S) \cap \{x \geq \lambda^{-\frac{1}{10}}\}$ and $R^s = N_h^{-n}(S) \cap \{y \geq \lambda^{-\frac{1}{10}}\}$.

Denote by R^* the intersection (see Fig. 4):

$$R^* = H^{k(h)} \circ \Psi_h^{-1} \circ C_h^{-1}(R^u) \cap \Psi_h^{-1} \circ C_h^{-1}(R^s).$$

Now consider the rectangles

(16) $$S_0 = S \cap N_h^{-1}(S) \quad \text{and} \quad S_1 = N_h^{-n} \circ C_h \circ \Psi_h \circ H^{-k(h)}(R^*)$$

and define the first return map

$$T(x,y) = \begin{cases} N_h(x,y), & \text{if } (x,y) \in S_0; \\ N_h^n \circ C_h \circ \Psi_h \circ H^{k(h)} \circ \Psi_h^{-1} \circ C_h^{-1} \circ N_h^n(x,y), & \text{if } (x,y) \in S_1. \end{cases}$$

5.2. Renormalization.

We are going to prove that the map T has a hyperbolic invariant set in S and to estimate its Hausdorff dimension with respect to the parameter h. It is convenient to renormalize the map T. Namely, using the approach from [Du3], set

$$\rho : S \to [0,2] \times [0,2], \ \rho(x,y) = (\Delta_h^n(xy)x, \Delta_h^n(xy)y),$$

and define

$$\rho(S_0) = \widetilde{S}_0, \ \rho(S_1) = \widetilde{S}_1, \ \widetilde{T} : \widetilde{S}_0 \cup \widetilde{S}_1 \to [0,2] \times [0,2], \ \widetilde{T} = \rho \circ T \circ \rho^{-1}.$$

Notice that ρ^{-1} is defined by

$$\rho^{-1}(x,y) = (\Delta_h^{-n}(t(xy))x, \Delta_h^{-n}(t(xy))y),$$

where $t(s)$ is such that $t(\Delta_h^{2n}(xy)xy) = xy$, or, equivalently,

(17) $$t(s) \cdot \Delta_h^{2n}(t(s)) = s.$$

Lemma 8. *For some $C > 0$ independent of h the following holds. If $(x_0, y_0) \in S$, $\rho(x_0, y_0) = (x, y)$, and $s = xy$ then $|t(s)|$, $|t'(s)|$, and $|t''(s)|$ are bounded by $C\lambda^{-2n}$.*

Proof of Lemma 8. Since $(x_0, y_0) \in S$, we have

$$0 \leq t(s) = t(\Delta_h^n(x_0y_0)x_0 \cdot \Delta_h^n(x_0y_0)y_0) = x_0y_0 \leq 4\lambda^{-2n}.$$

Differentiating (17) we get

$$t'(s)\Delta_h^{2n}(t(s)) + 2nt(s)\Delta_h^{2n-1}(t(s))\Delta_h'(t(s))t'(s) = 1,$$

therefore

$$|t'(s)| = \left| \Delta_h^{-2n}(t(s)) \frac{1}{1 + 2nt(s)\frac{\Delta_h'(t(s))}{\Delta_h(t(s))}} \right| \leq 4\lambda^{-2n}$$

Differentiating once again, we get

$$t''(s)(\lambda^{2n} + O(n)) + t'(s)O(n) = 0, \ \text{so} \ |t''(s)| \leq \frac{|t'(s)|O(n)}{\lambda^{2n}|1 + O(n\lambda^{-2n})|} < \lambda^{-2n}$$

$\square$

Notice that

$$\widetilde{T}|_{\widetilde{S}_0}(x,y) = \rho \circ N_h \circ \rho^{-1}(x,y) = (\Delta_h(t(xy))x, \Delta_h^{-1}(t(xy))y).$$

Lemma 9. *We have*

$$D\widetilde{T}|_{\widetilde{S}_0}(x,y) = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} + \begin{pmatrix} O(\lambda^{-2n}h) & O(\lambda^{-2n}h) \\ O(\lambda^{-2n}h) & O(\lambda^{-2n}h) \end{pmatrix}.$$

Proof of Lemma 9. Differentiating explicitly we get

$$DT\widetilde{\,}|_{\widetilde{S}_0}(x,y) = \begin{pmatrix} \Delta_h(t(xy)) + xy\Delta'_h(t(xy))t'(xy) & x^2\Delta'_h(t(xy))t'(xy) \\ -\frac{y^2}{\Delta_h^2(t(xy))}\Delta'_h(t(xy))t'(xy) & \Delta_h^{-1}(t(xy)) - xyt'(xy)\frac{\Delta'_h(t(xy))}{\Delta_h^2(t(xy))} \end{pmatrix}.$$

Now the required estimates follow from Lemmas 7 and 8. $\qquad\square$

In order to study $\widetilde{T}|_{\widetilde{S}_1}$ one can introduce the following maps:

$$G : \widetilde{S}_1 \to R^u, \ G = N_h^n \circ \rho^{-1}, \ G(x,y) = (x, \Delta_h^{-2n}(t(xy))y),$$

and

$$\widehat{G} : R^s \to \widetilde{S}_1, \ \widehat{G} = \rho \circ N_h^n, \ \widehat{G}(x,y) = (\Delta_h^{2n}(xy)x, y).$$

With this notation we have

$$\widetilde{T}|_{\widetilde{S}_1} = \widehat{G} \circ C_h \circ \Psi_h \circ H^{k(h)} \circ \Psi_h^{-1} \circ C_h^{-1} \circ G$$

5.3. Cone condition.
The coordinate changes $C_h \circ \Psi_h$ and $\Psi_h^{-1} \circ C_h^{-1}$ have uniformly bounded C^3-norms. Assume that their C^3-norms are bounded by some constant C_0.

Let us introduce the following cone fields in $\tilde{S}_0 \cup \tilde{S}_1$:

(18) $\qquad K^u(x,y) = \{\bar{v} = (v_1, v_2) \in T_{(x,y)}\tilde{S}_i \mid |v_1| > 0.01C_0^{-6}h^{-1-\nu}|v_2|\}, \ \ \text{and}$

(19) $\qquad K^s(x,y) = \{\bar{v} = (v_1, v_2) \in T_{(x,y)}\tilde{S}_i \mid |v_2| > 0.01C_0^{-6}h^{-1-\nu}|v_1|\}.$

Lemma 10. (Cone condition for $\tilde{S}_0$) *For small enough h the following holds.*

For every vector $\bar{v} \in K^u(x,y)$, $(x,y) \in \tilde{S}_0$, we have $D\tilde{T}_{(x,y)}(\bar{v}) \in K^u(\tilde{T}(x,y))$, and if $D\tilde{T}_{(x,y)}(\bar{v}) = \bar{w} \equiv (w_1, w_2)$ then $|w_1| \geq \lambda^{0.9}|v_1|$.

For every vector $\bar{v} \in K^s(x,y)$, $(x,y) \in \tilde{T}(\tilde{S}_0)$, we have $D\tilde{T}_{(x,y)}^{-1}(\bar{v}) \in K^s(\tilde{T}^{-1}(x,y))$, and if $D\tilde{T}_{(x,y)}^{-1}(\bar{v}) = \bar{w} \equiv (w_1, w_2)$ then $|w_2| \geq \lambda^{0.9}|v_2|$.

Proof of Lemma 10. This follows directly from Lemma 9. $\qquad\square$

Lemma 11. (Cone condition for $\tilde{S}_1$) *For small enough h the following holds.*

For every vector $\bar{v} \in K^u(x,y)$, $(x,y) \in \tilde{S}_1$, we have $D\tilde{T}_{(x,y)}(\bar{v}) \in K^u(\tilde{T}(x,y))$, and if $D\tilde{T}_{(x,y)}(\bar{v}) = \bar{w} \equiv (w_1, w_2)$ then $|w_1| \geq 0.01C_0^{-4}h^{-1-\nu}|v_1|$ and $|\bar{w}| \leq 25C_0^4h^{-1-\nu}|\bar{v}|$.

For every vector $\bar{v} \in K^s(x,y)$, $(x,y) \in \tilde{T}(\tilde{S}_1)$, we have $D\tilde{T}_{(x,y)}^{-1}(\bar{v}) \in K^s(\tilde{T}^{-1}(x,y))$, and if $D\tilde{T}_{(x,y)}^{-1}(\bar{v}) = \bar{w} \equiv (w_1, w_2)$ then $|w_2| \geq 0.01C_0^{-4}h^{-1-\nu}|v_2|$ and $|\bar{w}| \leq 25C_0^4h^{-1-\nu}|\bar{v}|$.

Before to begin the proof of Lemma 11 we will formulate and proof two extra lemmas that give estimates of the angle between images of vectors under linear maps.

Lemma 12. *For any two vectors $\bar{u}_1, \bar{u}_2$ and any linear map $A : \mathbb{R}^2 \to \mathbb{R}^2$ the following inequality holds:*

$$\sin \angle(A\bar{u}_1, A\bar{u}_2) \leq \|A\| \cdot \|A^{-1}\| \cdot |\sin \angle(\bar{u}_1, \bar{u}_2)|$$

Proof of Lemma 12. Take two vectors $\bar{s}_1$ and $\bar{s}_2$ such that $\bar{s}_2 \perp (\bar{s}_1 - \bar{s}_2)$ and $\bar{s}_1 \| \bar{u}_1$, $\bar{s}_2 \| \bar{u}_2$. In this case $|\sin \angle(\bar{u}_1, \bar{u}_2)| = \frac{|\bar{s}_1 - \bar{s}_2|}{|\bar{s}_1|}$. Now we have

$$\sin \angle(A\bar{u}_1, A\bar{u}_2) \leq \frac{|A\bar{s}_1 - A\bar{s}_2|}{|A\bar{s}_1|} \leq \frac{\|A\| |\bar{s}_1 - \bar{s}_2|}{\|A^{-1}\|^{-1}|\bar{s}_1|} = \|A\| \cdot \|A^{-1}\| \cdot |\sin \angle(\bar{u}_1, \bar{u}_2)|$$

$\square$

Lemma 13. *For any vector $\bar{u} \in \mathbb{R}^2$, $\bar{u} \neq 0$, and any linear maps $A, B : \mathbb{R}^2 \to \mathbb{R}^2$ the following inequality holds:*

$$\sin \angle(A\bar{u}, B\bar{u}) \leq \|A\| \cdot \|A - B\|.$$

Proof of Lemma 13.

$$\sin \angle(A\bar{u}, B\bar{u}) \leq \frac{|A\bar{u} - B\bar{u}|}{|A\bar{u}|} \leq \frac{\|A - B\|}{\|A\|^{-1}} = \|A\| \cdot \|A - B\|.$$

$\square$

Proof of Lemma 11. We will prove the first part of the statement. The proof of the second part is completely the same.

Take a vector $\bar{v} \equiv (v_1, v_2) \in K^u(x, y)$, $(x, y) \in \tilde{S}_1$.

Consider the following points:

$$P_1 = (x, y) \in \widetilde{S}_0, \quad P_2 = G(P_1) \in R^u, \quad P_3 = \Psi_h^{-1} \circ C_h^{-1}(P_2) \in \Pi_{r_0, E_0},$$

$$P_4 = H^{k(h)}(P_3) \in \Pi_{r_0, E_0}, \quad P_5 = C_h \circ \Psi_h(P_4) \in R^s, \quad P_6 = \widehat{G}(P_5) \in T(\widetilde{S}_1),$$

and denote by (x_i, y_i) the coordinates of the point P_i, $i = 1, \ldots, 6$. We will follow the image of the vector along this sequence of points and estimate the angle between that image and coordinate axes and the size of the image. Denote by $\bar{v}^{(i)} = (v_1^{(i)}, v_2^{(i)})$ the image of $\bar{v}$ at the point P_i, $i = 1, \ldots, 6$. In particular, $\bar{v}^{(1)} = \bar{v}$.

Step 1. Let us first estimate the norm and inclination of the vector $\bar{v}^{(2)} = DG(\bar{v}^{(1)})$. We have

$$DG(x, y) = \begin{pmatrix} 1 & 0 \\ \frac{2ny^2 \Delta_h'(t(xy))t'(xy)}{\Delta_h^{2n+1}(t(xy))} & \Delta_h^{-2n}(t) - \frac{2nxy\Delta_h'(t(xy))t'(xy)}{\Delta_h^{2n+1}(t(xy))} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ O(n\lambda^{-4n}h) & \lambda^{-2n} + O(n\lambda^{-4n}h) \end{pmatrix}.$$

Therefore

$$\bar{v}^{(2)} = DG(\bar{v}^{(1)}) = \begin{pmatrix} v_1 \\ v_1 O(n\lambda^{-4n}h) + v_2(\lambda^{-2n} + O(n\lambda^{-4n}h)) \end{pmatrix},$$

and hence (since $\bar{v} \equiv (v_1, v_2) \in K^u(x, y)$) we have $\frac{1}{2}|\bar{v}^{(1)}| \leq |\bar{v}^{(2)}| \leq 2|\bar{v}^{(1)}|$, and

$$\frac{|v_2^{(2)}|}{|v_1^{(2)}|} = O(n\lambda^{-4n}h) + \lambda^{-2n}\frac{v_2^{(1)}}{v_1^{(1)}} + \frac{v_2^{(1)}}{v_1^{(1)}}O(n\lambda^{-4n}h) < 200\lambda^{-2n}C_0^6 h^{1+\nu}$$

Step 2. We have $\bar{v}^{(3)} = D_{P_2}(\Psi_h^{-1} \circ C_h^{-1})\bar{v}^{(2)}$. Therefore $C_0^{-1}|\bar{v}^{(2)}| \leq |\bar{v}^{(3)}| \leq C_0|\bar{v}^{(2)}|$. Let us estimate the angle between $\bar{v}^{(3)}$ and the vector $\bar{e}_1 = (1, 0)$. Let P^* be a projection of the point P_2 to the line $\{y = 0\}$. Then $\mathrm{dist}(P_2, P^*) \leq 2\lambda^{-2n}$. Since the image of the line $\{y = 0\}$ under the map

$\Psi_h^{-1} \circ C_h^{-1}$ is a line $\{E = 0\}$, the image of the vector $\bar{e}_1 = (1, 0)$ under the differential $D(\Psi_h^{-1} \circ C_h^{-1})$ has the form $(s, 0) = s\bar{e}_1$. Now we have

$$\angle(\bar{v}^{(3)}, \bar{e}_1) = \angle(\bar{v}^{(3)}, s\bar{e}_1) = \angle\left(D_{P_2}(\Psi_h^{-1} \circ C_h^{-1})\bar{v}^{(2)}, D_{P^*}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_1\right) \le$$

$$\le \angle\left(D_{P_2}(\Psi_h^{-1} \circ C_h^{-1})\bar{v}^{(2)}, D_{P_2}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_1\right) + \angle\left(D_{P_2}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_1, D_{P^*}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_1\right)$$

Now let us estimate each of the summands. Since all the angles that we consider are small, we can always assume that $\alpha < 2\sin\alpha < 2\alpha$ for all angles α that we consider. Due to Lemma 12 we have

$$(20) \quad \angle\left(D_{P_2}(\Psi_h^{-1} \circ C_h^{-1})\bar{v}^{(2)}, D_{P_2}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_1\right) \le$$

$$\le 2\sin\angle\left(D_{P_2}(\Psi_h^{-1} \circ C_h^{-1})\bar{v}^{(2)}, D_{P_2}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_1\right) \le$$

$$\le 2C_0^2 |\sin\angle(\bar{v}^{(2)}, \bar{e}_1)| \le 2C_0^2 \cdot 200\lambda^{-2n} C_0^6 h^{1+\nu} = 400C_0^8 \lambda^{-2n} h^{1+\nu}$$

Due to Lemma 13 we have

$$(21) \quad \angle\left(D_{P_2}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_1, D_{P^*}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_1\right) \le$$

$$\le 2\sin\angle\left(D_{P_2}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_1, D_{P^*}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_1\right) \le$$

$$\le 2C_0 \cdot C_0 \operatorname{dist}(P_2, P^*) \le 4C_0^2 \lambda^{-2n}$$

Finally (if h is small enough and $\lambda = e^h$ is close to 1) we have

$$(22) \qquad\qquad \angle(\bar{v}^{(3)}, \bar{e}_1) \le 400C_0^8 \lambda^{-2n} h^{1+\nu} + 4C_0^2 \lambda^{-2n} < 5C_0^2 \lambda^{-2n}$$

Step 3. Since $H(t, E) = (t + h, E)$, the estimates for $\bar{v}^{(3)}$ work for $\bar{v}^{(4)} = DH^{k(h)}(\bar{v}^{(3)})$ also.

Step 4. Consider $\bar{v}^{(5)} = D_{P_4}(C_h \circ \Psi_h)\bar{v}^{(4)} \in T_{P_5}\widehat{U}$. Notice that

$$(23) \qquad\qquad |\bar{v}^{(5)}| \ge C_0^{-1}|\bar{v}^{(4)}| \ge C_0^{-2}|\bar{v}^{(2)}| > \frac{1}{2}C_0^{-2}|\bar{v}^{(1)}| \quad \text{and}$$

$$(24) \qquad\qquad |\bar{v}^{(5)}| \le C_0|\bar{v}^{(4)}| \le C_0^2|\bar{v}^{(2)}| \le 2C_0^2|\bar{v}^{(1)}|.$$

Now let us estimate the angle between $\bar{v}^{(5)}$ and the axis Oy. Let $P^{\#}$ be a projection of the point P_5 on the line $\{x = 0\}$. Take the vector $\bar{e}_2 = (0, 1) \in T_{P^{\#}}\widehat{U}$ and consider the image $D_{P^{\#}}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_2 \in T_{\Psi_h^{-1} \circ C_h^{-1}(P^{\#})}\Pi_{r_0, E_0}$. The vector $D_{P^{\#}}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_2$ is tangent to the graph of the function $\Theta(t)$, and due to (10)

$$(25) \qquad\qquad \frac{1}{2}\mu(h) < \angle(D_{P^{\#}}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_2, \bar{e}_1) < 2\mu(h).$$

From (22) we have

$$\frac{1}{5}\mu(h) < \angle(D_{P^{\#}}(\Psi_h^{-1} \circ C_h^{-1})\bar{e}_2, \bar{v}^{(4)}) < 5\mu(h).$$

Notice that $\operatorname{dist}(\Psi_h^{-1} \circ C_h^{-1}(P^{\#}), P_4) \le 2C_0\lambda^{-2n} \le 4C_0\mu(h)h^{1+\nu}$. This implies (in the way similar to Step 3) that for small enough h

$$(26) \qquad\qquad \angle(\bar{v}^{(5)}, \bar{e}_2) < 5C_0^2\mu(h) + C_0 \cdot C_0 \cdot 4C_0\mu(h)h^{1+\nu} < 6C_0^2\mu(h),$$

$$(27) \qquad\qquad \angle(\bar{v}^{(5)}, \bar{e}_2) > \frac{1}{5}\mu(h)C_0^{-2} - 4C_0^3\mu(h)h^{1+\nu} > \frac{1}{6}C_0^{-2}\mu(h).$$

Step 5. Since $\widehat{G}(x,y) = (\Delta_h^{2n}(xy)x, y)$, we have

$$(28) \quad D\widehat{G}(x,y) = \begin{pmatrix} \Delta_h^{2n}(xy) + 2nxy\Delta_h^{2n-1}(xy)\Delta_h'(xy) & 2nx^2\Delta_h^{2n-1}(xy)\Delta_h'(xy) \\ 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda^{2n}(1 + O(n\lambda^{-2n}h)) & O(n\lambda^{-2n}h)) \\ 0 & 1 \end{pmatrix},$$

hence

$$D\widehat{G}(\bar{v}^{(5)}) = \begin{pmatrix} v_1^{(6)} \\ v_2^{(6)} \end{pmatrix} = \begin{pmatrix} \lambda^{2n}v_1^{(5)}(1 + O(n\lambda^{-2n}h)) + v_2^{(5)}O(n\lambda^{-2n}h)) \\ v_2^{(5)} \end{pmatrix}.$$

Therefore $|v_2^{(6)}| = |\bar{v}^{(5)}| \leq 2C_0^2|\bar{v}^{(1)}|$ and

$$(29) \quad |v_1^{(6)}| \geq |v_2^{(5)}| \cdot \left| \lambda^{2n}\frac{v_1^{(5)}}{v_2^{(5)}}(1 + O(n\lambda^{-2n}h)) + O(n\lambda^{-2n}h) \right| \geq$$

$$\geq \frac{1}{2}|\bar{v}^{(5)}|\left|\frac{1}{6}C_0^{-2}h^{-1-\nu}(1 + O(n\lambda^{-2n}h)) + O(n\lambda^{-2n}h)\right| \geq \frac{1}{20}C_0^{-2}h^{-1-\nu}|\bar{v}^{(5)}| \geq \frac{1}{40}C_0^{-4}h^{-1-\nu}|\bar{v}^{(1)}|$$

This implies that $\frac{|v_1^{(6)}|}{|v_2^{(6)}|} > \frac{1}{80}C_0^{-6}h^{-1-\nu}$, and hence $\bar{v}^{(6)} \in K^u(P_6)$.

Also we have

$$(30) \quad |\bar{v}^{(6)}| \leq |v_1^{(6)}| + |v_2^{(6)}| \leq |\bar{v}^{(5)}| + 2|\bar{v}^{(5)}|\left|\lambda^{2n}6C_0^2\mu(h)(1 + O(n\lambda^{-2n}h)) + O(n\lambda^{-2n}h)\right| \leq$$

$$\leq 2C_0^2|\bar{v}^{(1)}| + 2 \cdot 2C_0^2|\bar{v}^{(1)}|\left|\lambda^{2n}6C_0^2\mu(h)(1 + O(n\lambda^{-2n}h)) + O(n\lambda^{-2n}h)\right| \leq 25C_0^4h^{-1-\nu}|\bar{v}^{(1)}|$$

$$\square$$

5.4. **Markov partition and its thickness.** Standard arguments of the hyperbolic theory (see, for example, [IL]) show that the Cone condition (Lemmas 10 and 11) together with the geometry of the map $\tilde{T}$ imply the existence of the hyperbolic fixed point $\mathbf{Q}$ of the map $\tilde{T}$ in $\tilde{S}_1 \cap \tilde{T}(\tilde{S}_1)$. Our choice of the homoclinic points q_h^u and q_h^s implies that the eigenvalues of $\mathbf{Q}$ are positive. Denote the heteroclinic point where $W_{loc}^s(\mathbf{Q})$ intersects $W^u(O) = \{(0,x)|x \in \mathbb{R}\}$ by $(x_s, 0)$, and the heteroclinic point where $W_{loc}^u(\mathbf{Q})$ intersects $W^s(O) = \{(y,0)|y \in \mathbb{R}\}$ by $(0, y_u)$.

Denote the segments of stable and unstable manifolds that connect the fixed points O and $\mathbf{Q}$ with these heteroclinic points by

$\gamma^u(O)$ – connects O and $(x_s, 0)$, $\quad \gamma^s(O)$ – connects O and $(0, y_u)$;

$\gamma^u(\mathbf{Q})$ – connects $\mathbf{Q}$ and $(0, y_u)$, $\quad \gamma^s(\mathbf{Q})$ – connects $\mathbf{Q}$ and $(x_s, 0)$.

Notice that $\gamma^s(\mathbf{Q}) \subset \tilde{S}_1$ and $\gamma^u(\mathbf{Q}) \subset \tilde{T}(\tilde{S}_1)$.

Let $\mathbf{S}$ be the square formed by $\gamma^u(O)$, $\gamma^s(O)$, $\gamma^u(\mathbf{Q})$ and $\gamma^s(\mathbf{Q})$, $\mathbf{S} \subset \tilde{S}$.

Now define $\mathbf{S}_0 = \rho(\rho^{-1}(\mathbf{S}) \cap N_h^{-1} \circ \rho^{-1}(\mathbf{S})) \subset \tilde{S}_0$ and $\mathbf{S}_1 = \rho(S_1 \cap \rho^{-1}(\mathbf{S})) \subset \tilde{S}_1$, see Fig 5. Notice that one of the vertical edges of $\mathbf{S}_1$ is $\gamma^s(\mathbf{Q})$ and another is an intersection of $\mathbf{S}$ and a vertical edge of $\tilde{S}_1$, and therefore it intersects $W^u(O)$ at the point $\rho(N_h^{-n}(1,0)) = \rho((\lambda^{-n}, 0)) = (1, 0)$. Similarly, $\tilde{T}(\mathbf{S}_1)$ has a vertical edge $[1, y_u] \subset Oy$.
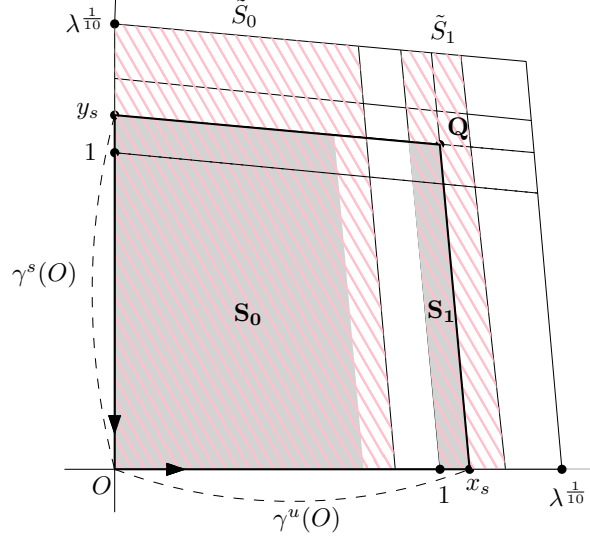
FIGURE 5. Rectangles $\tilde{S}_0$, $\tilde{S}_1$ and $\mathbf{S_0}$, $\mathbf{S_1}$.

Define now $\mathbf{T} = \tilde{T}|_{\mathbf{S}}$. The maximal invariant set of $\mathbf{T}$ in $\mathbf{S}$, $\Lambda = \cap_{n\in\mathbb{Z}}\mathbf{T}^{-n}(\mathbf{S})$, is a "horseshoe"-type basic set with Markov partition $\mathcal{P} = \{\mathbf{S}_0, \mathbf{S}_1\}$. The map $\mathbf{T} : \mathbf{S}_0 \cup \mathbf{S}_1 \to \mathbf{S}$ belongs to class $\mathcal{F}$ (see definition 3).

Consider now the Markov partitions

$$\mathcal{P}^s = \{[0, \lambda^{-1}x_s], [1, x_s]\} \quad \text{and} \quad \mathcal{P}^u = \{[0, \lambda^{-1}y_u], [1, y_u]\}$$

of the Cantor sets $K^s \subset Ox$ and $K^u \subset Oy$ associated with the hyperbolic set Λ. We have

$$\tau_L(\mathcal{P}^s) = \frac{\lambda^{-1}x_s}{1 - \lambda^{-1}x_s}, \quad \tau_R(\mathcal{P}^s) = \frac{x_s - 1}{1 - \lambda^{-1}x_s},$$

$$\tau_L(\mathcal{P}^u) = \frac{\lambda^{-1}y_u}{1 - \lambda^{-1}y_u}, \quad \tau_R(\mathcal{P}^u) = \frac{y_u - 1}{1 - \lambda^{-1}y_u}.$$

Lemma 14. *The following estimates hold for all $h \in (0, h_0)$ if h_0 is small enough:*

$$\frac{1}{2}h^{-1} \leq \tau_L(\mathcal{P}^s) \leq 2h^{-1}, \quad 0.01C_0^{-4}h^\nu \leq \tau_R(\mathcal{P}^s) \leq 250C_0^4h^\nu,$$

$$\frac{1}{2}h^{-1} \leq \tau_L(\mathcal{P}^u) \leq 2h^{-1}, \quad 0.01C_0^{-4}h^\nu \leq \tau_R(\mathcal{P}^u) \leq 250C_0^4h^\nu.$$

Proof of Lemma 14. We will prove only estimates for the partition $\mathcal{P}^s$ (for $\mathcal{P}^u$ everything is the same). Notice first that $1 - \lambda^{-1}x_s \leq x_s(1 - \lambda^{-1})$. Therefore

$$\tau_L(\mathcal{P}^s) = \frac{\lambda^{-1}x_s}{1 - \lambda^{-1}x_s} \geq \frac{\lambda^{-1}x_s}{x_s - \lambda^{-1}x_s} = \frac{\lambda^{-1}}{1 - \lambda^{-1}} = \frac{1}{\lambda - 1} = \frac{1}{e^h - 1} \geq \frac{1}{2}h^{-1},$$

since $e^h - 1 \leq 2h$ for small h.

On the other hand, since $x_s \in (1, \lambda^{\frac{1}{10}})$

$$\tau_L(\mathcal{P}^s) = \frac{\lambda^{-1}x_s}{1 - \lambda^{-1}x_s} \leq \frac{\lambda^{-1}\lambda^{\frac{1}{10}}}{1 - \lambda^{-1}\lambda^{\frac{1}{10}}} = \frac{1}{\lambda^{\frac{9}{10}} - 1} = \frac{1}{e^{\frac{9}{10}h} - 1} \leq \frac{10}{9}h^{-1} < 2h^{-1}$$

Now let us estimate $\tau_R(\mathcal{P}^s)$. Denote by I the segment of $W^u_{loc}(O)$ between the points $(1,0)$ and $(x_s, 0)$ (i.e. the bottom horizontal edge of $\mathbf{S}_1$), $|I| = x_s - 1$. Due to Lemma 11

$$0.01 C_0^{-4} h^{-1-\nu} |I| \leq |\mathbf{T}(I)| \leq 25 C_0^4 h^{-1-\nu} |I|.$$

Since $\frac{1}{2} \leq |\mathbf{T}(I)| \leq 2$ (this follows from the Cone condition again), we have

$$\frac{1}{50} C_0^{-4} h^{1+\nu} \leq |I| \leq 200 C_0^4 h^{1+\nu}$$

Hence

$$\tau_R(\mathcal{P}^s) \geq \frac{\frac{1}{50} C_0^{-4} h^{1+\nu}}{1 - \lambda^{-1} x_s} \geq \frac{\frac{1}{50} C_0^{-4} h^{1+\nu}}{1 - e^{-h}} \geq 0.01 C_0^{-4} h^{\nu}$$

since $1 - e^{-h} \leq h$ for $h \geq 0$, and

$$\tau_R(\mathcal{P}^s) \leq \frac{200 C_0^4 h^{1+\nu}}{1 - \lambda^{-1} x_s} \leq \frac{200 C_0^4 h^{1+\nu}}{1 - \lambda^{-1} \lambda^{\frac{1}{10}}} \leq \frac{200 C_0^4 h^{1+\nu}}{1 - \lambda^{-\frac{9}{10}}} = \frac{200 C_0^4 h^{1+\nu}}{1 - e^{-\frac{9}{10}h}} \leq 250 C_0^4 h^{\nu}.$$

$\square$

Lemma 15. *If $h_0 > 0$ is small enough then for all $h \in (0, h_0)$ we have*

$$\mathrm{dist}(\mathbf{S}_0, \mathbf{S}_1) \geq 0.1 h \quad \text{and} \quad \mathrm{dist}(\mathbf{T}(\mathbf{S}_0), \mathbf{T}(\mathbf{S}_1)) \geq 0.1 h.$$

Proof of Lemma 15. Notice that the vertical boundaries of $\mathbf{S}_0$ and $\mathbf{S}_1$ are tangent to the cone field $\{K^u\}$. Consider the left vertical edge of $\mathbf{S}_1$ and the right vertical edge of $\mathbf{S}_0$. Their lowest points are $(1, 0)$ and $(\lambda^{-1} x_s, 0)$, and the distance between them is equal to

$$1 - \lambda^{-1} x_s \geq 1 - \lambda^{-1} \lambda^{\frac{1}{10}} = 1 - \lambda^{-\frac{9}{10}} = 1 - e^{-\frac{9}{10}h} \geq \frac{9}{20} h$$

if $h \in (0, h_0)$ and h_0 is small enough. From the cone condition we have that the difference between x-coordinates of any two points on those edges is greater than

$$\frac{9}{20} h - 2 \cdot 100 C_0^6 h^{1+\nu} \geq \frac{9}{20} h \left(1 - \frac{4000}{9} C_0^6 h^{\nu}\right) \geq \frac{9}{40} h > 0.1 h$$

if h is small enough.

The proof of the second inequality is completely similar. $\square$

5.5. Estimates of derivatives: verification of the conditions of Distortion Theorem.

We proved that the map $\mathbf{T} : \mathbf{S}_0 \cup \mathbf{S}_1 \to \mathbf{S}$ has an invariant locally maximal hyperbolic set Λ which is a two-component Smale horseshoe (i.e. $\mathbf{T}$ belongs to the class $\mathcal{F}$) and obtained estimates of the lateral thicknesses of the corresponding Markov partitions. In order to get estimates of the lateral thicknesses of the related Cantor sets we need to estimate the distortion of the corresponding mappings.

Denote the differential of the map $\mathbf{T} : \mathbf{S}_0 \cup \mathbf{S}_1 \to \mathbf{S}$ by $D\mathbf{T} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where a, b, c and d are smooth functions over $\mathbf{S}_0 \cup \mathbf{S}_1$. Then the differential of the inverse map $\mathbf{T}^{-1} : \mathbf{T}(\mathbf{S}_0) \cup \mathbf{T}(\mathbf{S}_1) \to \mathbf{S}$ has the form $D\mathbf{T}^{-1} = \begin{pmatrix} \tilde{d} & -\tilde{b} \\ -\tilde{c} & \tilde{a} \end{pmatrix}$, where $\tilde{a} = a \circ \mathbf{T}^{-1}$, $\tilde{b} = b \circ \mathbf{T}^{-1}$, $\tilde{c} = c \circ \mathbf{T}^{-1}$ and $\tilde{d} = d \circ \mathbf{T}^{-1}$. Notice that this notation agrees with the notation of Definition 6.

Lemma 16. *Consider the restriction of the map* $\mathbf{T}$ *to the rectangle* $\mathbf{S}_0$*. There exists a constant* $C_1 > 1$ *(independent of* h*) such that*

(1) $\lambda(1 - C_1 h \lambda^{-2n}) \leq |a| \leq \lambda(1 + C_1 h \lambda^{-2n})$,

(2) $|d| \leq \lambda^{-1}(1 + C_1 h \lambda^{-2n})$,

(3) $|b|, |c| \leq C_1 h \lambda^{-2n}$,

(4) $\left|\frac{\partial a}{\partial y}\right|, \left|\frac{\partial a}{\partial x}\right|, \left|\frac{\partial b}{\partial x}\right|, \left|\frac{\partial b}{\partial y}\right|, \left|\frac{\partial c}{\partial x}\right|, \left|\frac{\partial c}{\partial y}\right|, \left|\frac{\partial d}{\partial y}\right|, \left|\frac{\partial d}{\partial x}\right| \leq C_1 h \lambda^{-2n}$,

(5) $\left|\frac{\partial \tilde{a}}{\partial y}\right|, \left|\frac{\partial \tilde{a}}{\partial x}\right|, \left|\frac{\partial \tilde{b}}{\partial x}\right|, \left|\frac{\partial \tilde{b}}{\partial y}\right|, \left|\frac{\partial \tilde{c}}{\partial x}\right|, \left|\frac{\partial \tilde{c}}{\partial y}\right|, \left|\frac{\partial \tilde{d}}{\partial y}\right|, \left|\frac{\partial \tilde{d}}{\partial x}\right| \leq C_1 h \lambda^{-2n}$,

Proof of Lemma 16. Notice that Lemma 16 is a stronger version of Lemma 9. Since

$$\mathbf{T}|_{\mathbf{S}_0}(x,y) = \rho \circ N_h \circ \rho^{-1}(x,y) = (\Delta_h(t(xy))x, \Delta_h^{-1}(t(xy))y),$$

we have

$$a(x,y) = \Delta_h(t(xy)) + xy\Delta_h'(t(xy))t'(xy)$$
$$b(x,y) = x^2\Delta_h'(t(xy))t'(xy)$$

(31)
$$c(x,y) = -\frac{y^2}{\Delta_h^2(t(xy))}\Delta_h'(t(xy))t'(xy)$$

$$d(x,y) = \Delta_h^{-1}(t(xy)) - xyt'(xy)\frac{\Delta_h'(t(xy))}{\Delta_h^2(t(xy))}$$

Now the required estimated follow by a direct calculation from Lemmas 7 and 8. □

Lemma 17. *The variation of* $\log|a(x,y)|$ *in* $\mathbf{S}_0$ *is less than* $4C_1 h \lambda^{-2n}$*.*

Proof of Lemma 17. This follows immediately from the estimate (1) in Lemma 16. □

Lemma 18. *Consider the restriction of the map* $\mathbf{T}$ *to the rectangle* $\mathbf{S}_1$*. There exists a constant* $C_1 > 1$ *(independent of* h*) such that*

(1) $|d| \leq \lambda^{-2n}C_1$, $|b|, |c| \leq C_1$,

(2) $C_1^{-1} h^{-1-\nu} \leq |a| \leq C_1 h^{-1-\nu}$,

(3) $\left|\frac{\partial b}{\partial x}\right|, \left|\frac{\partial b}{\partial y}\right|, \left|\frac{\partial c}{\partial x}\right|, \left|\frac{\partial c}{\partial y}\right|, \left|\frac{\partial a}{\partial y}\right| \leq C_1$,

(4) $\left|\frac{\partial d}{\partial x}\right|, \left|\frac{\partial d}{\partial y}\right| \leq \lambda^{-2n}C_1$,

(5) $\left|\frac{\partial a}{\partial x}\right| \leq C_1 h^{-2-\nu}$,

(6) $\left|\frac{\partial \tilde{b}}{\partial x}\right|, \left|\frac{\partial \tilde{b}}{\partial y}\right|, \left|\frac{\partial \tilde{c}}{\partial x}\right|, \left|\frac{\partial \tilde{c}}{\partial y}\right|, \left|\frac{\partial \tilde{a}}{\partial x}\right| \leq C_1$,

(7) $\left|\frac{\partial \tilde{d}}{\partial y}\right|, \left|\frac{\partial \tilde{d}}{\partial x}\right| \leq \lambda^{-2n}C_1$,

(8) $\left|\frac{\partial \tilde{a}}{\partial y}\right| \leq C_1 h^{-2-\nu}$.

Proof of Lemma 18. We need to study the differential of the map

$$\mathbf{T}|_{\mathbf{S}_1} = \widehat{G} \circ C_h \circ \Psi_h \circ H^{k(h)} \circ \Psi_h^{-1} \circ C_h^{-1} \circ G, \quad \text{where} \quad G : \mathbf{S}_1 \to R^u, \ \widehat{G} : R^s \to \mathbf{S}_1.$$

In order to distinguish the points from R^u and from R^s let us denote the coordinates in R^u by $(\mathbf{x}, \mathbf{y})$ and the coordinates in R^s by $(\mathbf{X}, \mathbf{Y})$. Then the map $C_h \circ \Psi_h \circ H^{k(h)} \circ \Psi_h^{-1} \circ C_h^{-1} : R^u \to U$ can be represented as a composition

$$(\mathbf{x}, \mathbf{y}) \mapsto (t(\mathbf{x}, \mathbf{y}), E(\mathbf{x}, \mathbf{y})) \mapsto (\mathbf{X}(t, E), \mathbf{Y}(t, E)), \quad (t, E) \in \Pi_{r_0, E_0},$$

where

$$(t(\mathbf{x}, \mathbf{y}), E(\mathbf{x}, \mathbf{y})) = H^{k(h)} \circ \Psi_h^{-1} \circ C_h^{-1}(\mathbf{x}, \mathbf{y}) \quad \text{and} \quad (\mathbf{X}(t, E), \mathbf{Y}(t, E)) = C_h \circ \Psi_h(t, E).$$

Let us denote

$$\begin{pmatrix} a_0 & b_0 \\ c_0 & d_0 \end{pmatrix} = D(C_h \circ \Psi_h \circ H^{h(h)} \circ \Psi_h^{-1} \circ C_h^{-1}),$$

then

$$D\mathbf{T}|_{\mathbf{S}_1}(x, y) = \begin{pmatrix} a(x, y) & b(x, y) \\ c(x, y) & d(x, y) \end{pmatrix} = D\widehat{G}(\mathbf{X}, \mathbf{Y}) \begin{pmatrix} a_0(\mathbf{x}, \mathbf{y}) & b_0(\mathbf{x}, \mathbf{y}) \\ c_0(\mathbf{x}, \mathbf{y}) & d_0(\mathbf{x}, \mathbf{y}) \end{pmatrix} DG(x, y),$$

where

$$DG(x, y) = \begin{pmatrix} 1 & 0 \\ \frac{2ny^2 \Delta_h'(t(xy)) t'(xy)}{\Delta_h^{2n+1}(t(xy))} & \Delta_h^{-2n}(t) - \frac{2nxy \Delta_h'(t(xy)) t'(xy)}{\Delta_h^{2n+1}(t(xy))} \end{pmatrix},$$

$$D\widehat{G}(\mathbf{X}, \mathbf{Y}) = \begin{pmatrix} \Delta_h^{2n}(\mathbf{XY}) + 2n\mathbf{XY}\Delta_h^{2n-1}(\mathbf{XY})\Delta_h'(\mathbf{XY}) & 2n\mathbf{X}^2\Delta_h^{2n-1}(\mathbf{XY})\Delta_h'(\mathbf{XY}) \\ 0 & 1 \end{pmatrix}.$$

Therefore

$$a(x, y) = (\Delta_h^{2n}(\mathbf{XY}) + 2n\mathbf{XY}\Delta_h^{2n-1}(\mathbf{XY})\Delta_h'(\mathbf{XY})) \left(a_0(\mathbf{x}, \mathbf{y}) + \frac{2b_0(\mathbf{x}, \mathbf{y})ny^2\Delta_h'(t(xy))t'(xy)}{\Delta_h^{2n+1}(t(xy))} \right) +$$

$$+ 2n\mathbf{X}^2\Delta_h^{2n-1}(\mathbf{XY})\Delta_h'(\mathbf{XY}) \left(c_0(\mathbf{x}, \mathbf{y}) + \frac{2d_0(\mathbf{x}, \mathbf{y})ny^2\Delta_h'(t(xy))t'(xy)}{\Delta_h^{2n+1}(t(xy))} \right),$$

$$b(x, y) = (\Delta_h^{2n}(\mathbf{XY}) + 2n\mathbf{XY}\Delta_h^{2n-1}(\mathbf{XY})\Delta_h'(\mathbf{XY})) \left(\Delta_h^{-2n}(t(xy))b_0(\mathbf{x}, \mathbf{y}) - \frac{2b_0(\mathbf{x}, \mathbf{y})nxy\Delta_h'(t(xy))t'(xy)}{\Delta_h^{2n+1}(t(xy))} \right)$$

$$+ 2n\mathbf{X}^2\Delta_h^{2n-1}(\mathbf{XY})\Delta_h'(\mathbf{XY}) \left(\Delta_h^{-2n}(t(xy))d_0(\mathbf{x}, \mathbf{y}) - \frac{2d_0(\mathbf{x}, \mathbf{y})nxy\Delta_h'(t(xy))t'(xy)}{\Delta_h^{2n+1}(t(xy))} \right),$$

$$c(x, y) = c_0(\mathbf{x}, \mathbf{y}) + \frac{2d_0(\mathbf{x}, \mathbf{y})ny^2\Delta_h'(t(xy))t'(xy)}{\Delta_h^{2n+1}(t(xy))},$$

$$d(x, y) = \Delta_h^{-2n}(t(xy))d_0(\mathbf{x}, \mathbf{y}) - \frac{2d_0(\mathbf{x}, \mathbf{y})nxy\Delta_h'(t(xy))t'(xy)}{\Delta_h^{2n+1}(t(xy))}.$$

The C^3-norm of the map $C_h \circ \Psi_h \circ H^{k(h)} \circ \Psi_h^{-1} \circ C_h^{-1}$ (as well as of its inverse) is uniformly bounded by some constant independent of h. Together with Lemmas 7 and 8 this implies that $d(x, y) = O(\lambda^{-2n})$, $b(x, y) = O(1), c(x, y) = O(1)$, and this proves the statement (1). Inequality (2) follows directly from Lemma 11 for large enough C_1, for example, $C_1 > 100C_0^4$. Using the estimates

$$\mathbf{X}, \mathbf{y}, \frac{\partial \mathbf{X}}{\partial y}, \frac{\partial \mathbf{Y}}{\partial y} = O(\lambda^{-2n}), \quad \frac{\partial(\mathbf{XY})}{\partial x}, \frac{\partial \mathbf{X}}{\partial x} = O(1), \quad \frac{\partial(\mathbf{XY})}{\partial y} = O(\lambda^{-2n}),$$

$$\frac{\partial \mathbf{x}}{\partial x} = 1, \quad \frac{\partial \mathbf{x}}{\partial y} = 0, \quad \frac{\partial \mathbf{y}}{\partial x} = O(\lambda^{-4n}), \quad \frac{\partial \mathbf{y}}{\partial y} = O(\lambda^{-2n}), \ n = O(h^{-2}),$$

one can show that partial derivatives of $b(x, y)$ and $c(x, y)$ are bounded, and that $\frac{\partial d}{\partial x}, \frac{\partial d}{\partial y}$ are of order $O(\lambda^{-2n})$ (we omit the details of these tedious but straightforward calculations).

In order to study the partial derivatives of $a(x, y)$, let us first obtain some estimates for $a_0(\mathbf{x}, \mathbf{y})$. We have
$$a_0(\mathbf{x}, \mathbf{y}) = \frac{\partial}{\partial \mathbf{x}} \mathbf{X}(t(\mathbf{x}, \mathbf{y}), E(\mathbf{x}, \mathbf{y})) \quad \text{and} \quad \frac{\partial a_0}{\partial \mathbf{x}} = \frac{\partial^2}{\partial \mathbf{x}^2} \mathbf{X}(t(\mathbf{x}, \mathbf{y}), E(\mathbf{x}, \mathbf{y})).$$
In particular, since $E(\mathbf{x}, 0) = 0$,

$$(32) \quad \frac{\partial a_0}{\partial \mathbf{x}}(\mathbf{x}, 0) = \frac{d^2}{d\mathbf{x}^2} \mathbf{X}(t(\mathbf{x}, 0), 0) = \frac{d}{d\mathbf{x}} \left(\frac{\partial \mathbf{X}}{\partial t}(t(\mathbf{x}, 0), 0) \cdot \frac{\partial t}{\partial \mathbf{x}}(\mathbf{x}, 0) \right) =$$

$$= \frac{\partial^2 \mathbf{X}}{\partial t^2}(t(\mathbf{x}, 0), 0) \cdot \left(\frac{\partial t}{\partial \mathbf{x}}(\mathbf{x}, 0) \right)^2 + \frac{\partial \mathbf{X}}{\partial t}(t(\mathbf{x}, 0), 0) \cdot \frac{\partial^2 t}{\partial \mathbf{x}^2}(\mathbf{x}, 0) =$$

$$= \frac{\partial^2 \mathbf{X}}{\partial t^2}(t(\mathbf{x}, 0), 0) \cdot \left(\frac{\partial t}{\partial \mathbf{x}}(\mathbf{x}, 0) \right)^2 + a_0(\mathbf{x}, 0) \cdot \left(\frac{\partial t}{\partial \mathbf{x}}(\mathbf{x}, 0) \right)^{-1} \cdot \frac{\partial^2 t}{\partial \mathbf{x}^2}(\mathbf{x}, 0)$$

From the Cone condition (more precisely, from Steps 3-5 of the proof of Lemma 11) we know that $a_0(\mathbf{x}, 0) = O(\mu(h))$. Also since C^3-norms of maps $C_h \circ \Psi_h$ and $H^{k(h)} \circ \Psi_h^{-1} \circ C_h^{-1}$ are bounded by C_0, we have
$$\left| \frac{\partial a_0}{\partial \mathbf{x}}(\mathbf{x}, 0) \right| \leq O \left(\frac{\partial^2 \mathbf{X}}{\partial t^2}(t(\mathbf{x}, 0), 0) \right) + O(\mu(h)).$$
Now we need to estimate $\left| \frac{\partial^2 \mathbf{X}}{\partial t^2}(t(\mathbf{x}, 0), 0) \right|$. Notice that the image of the Oy ax under the map $C_h \circ \Psi_h$ is a graph of the function $E = \Theta(t)$, and therefore $\mathbf{X}(t, \Theta(t)) = 0$. This implies that
$$\frac{d}{dt}(\mathbf{X}(t, \Theta(t)) = 0 = \frac{\partial \mathbf{X}}{\partial t}(t, \Theta(t)) + \frac{\partial \mathbf{X}}{\partial E}(t, \Theta(t)) \cdot \dot{\Theta}(t)$$
and

$$(33) \quad \frac{d^2}{dt^2}(\mathbf{X}(t, \Theta(t)) = 0 =$$

$$= \frac{\partial^2 \mathbf{X}}{\partial t^2}(t, \Theta(t)) + 2\frac{\partial^2 \mathbf{X}}{\partial t \partial E}(t, \Theta(t))\dot{\Theta}(t) + \frac{\partial^2 \mathbf{X}}{\partial E^2}(t, \Theta(t))(\dot{\Theta}(t))^2 + \frac{\partial \mathbf{X}}{\partial E}(t, \Theta(t))\ddot{\Theta}(t)$$

Since for $\Theta(t)$, $\dot{\Theta}(t)$ and $\ddot{\Theta}(t)$ we have asymptotics (10) (see Theorem 7 and Remark 6), we get
$$\left| \frac{\partial^2 \mathbf{X}}{\partial t^2}(t, \Theta(t)) \right| \leq 2C_0 \cdot 2\mu(h) + C_0(2\mu(h))^2 + C_0 \cdot 4\pi h^{-1}\mu(h) = O(h^{-1}\mu(h))$$
if h is small enough.

At the same time by the mean value theorem we have
$$\left| \frac{\partial^2 \mathbf{X}}{\partial t^2}(t, 0) \right| \leq \left| \frac{\partial^2 \mathbf{X}}{\partial t^2}(t, \Theta(t)) \right| + C_0|\Theta(t)| < 20C_0 h^{-1}\mu(h) + C_0\frac{1}{\pi}h\mu(h) = O(h^{-1}\mu(h))$$
Finally we have
$$\frac{\partial a_0}{\partial \mathbf{x}}(\mathbf{x}, 0) = O(h^{-1}\mu(h)) + O(\mu(h)) = O(h^{-1}\mu(h)),$$
and since $\mathbf{y} = O(\lambda^{-2n}) = O(h^{1+\nu}\mu(h))$, by the mean value theorem we also have $\frac{\partial a_0}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{y}) = O(h^{-1}\mu(h))$. Now we have
$$\frac{\partial a}{\partial y}(x, y) = a_0(\mathbf{x}, \mathbf{y})2n\Delta_h^{2n-1}(\mathbf{XY})\Delta_h'(\mathbf{XY})\frac{\partial(\mathbf{XY})}{\partial y} + \frac{\partial a_0}{\partial y}(\mathbf{x}, \mathbf{y})\Delta_h^{2n}(\mathbf{XY}) + O(nh\lambda^{-2n}),$$

and since $a_0(\mathbf{x}, \mathbf{y}) = O(\mu(h))$ and $\frac{\partial a_0}{\partial y}(\mathbf{x}, \mathbf{y}) = O(\lambda^{-2n})$, we also have

$$\frac{\partial a}{\partial y}(x, y) = O(n\mu(h)\lambda^{2n}h\lambda^{-2n}) + O(1) + O(nh\lambda^{-2n}) = O(1).$$

Let us now estimate $\frac{\partial a}{\partial x}(x, y)$. We have

$$\frac{\partial a}{\partial x}(x, y) = \frac{\partial a_0}{\partial x}(\mathbf{x}, \mathbf{y})(\Delta_h^{2n}(\mathbf{XY}) + 2n\mathbf{XY}\Delta_h^{2n-1}(\mathbf{XY})\Delta_h'(\mathbf{XY})) +$$
$$+ a_0(\mathbf{x}, \mathbf{y})2n\frac{\partial(\mathbf{XY})}{\partial x}\Delta_h^{2n-1}(\mathbf{XY})\Delta_h'(\mathbf{XY}) + O(nh).$$

Therefore

$$\frac{\partial a}{\partial x}(x, y) = O(h^{-1}\mu(h))\left(O(\lambda^{2n}) + O(n\lambda^{-2n}\lambda^{2n}h)\right) + O(\mu(h)n\lambda^{2n}h) + O(nh) =$$
$$= O(h^{-1}h^{-1-\nu}) + O(h^{-1-\nu}nh) + O(nh) = O(h^{-2-\nu}).$$

This proves statements (3), (4), and (5). Estimates (6), (7), and (8) are symmetric to the estimates (3), (4), and (5). □

Lemma 19. *The variation of $\log|a(x, y)|$ in $\mathbf{S}_1$ is less than $600C_1^2C_0^6h^\nu$.*

Proof of Lemma 19. Take two points (x_1, y_1) and (x_2, y_2) from $\mathbf{S}_1$. We want to estimate $|\log a(x_1, y_1) - \log a(x_2, y_2)|$ by using the mean value theorem. Generally speaking, the set $\mathbf{S}_1$ is not convex, so we need some preparations to apply it.

Let $\tilde{\gamma}$ be the intersection $\tilde{\gamma} = \mathbf{T}(\mathbf{S}_1) \cap \{x = \frac{1}{2}\}$. Then $\hat{\gamma} = \mathbf{T}^{-1}(\tilde{\gamma})$ is a smooth curve tangent to the cone field $\{K^s\}$, $\hat{\gamma} \subset \mathbf{S}_1$. Denote $\hat{x}_1 = \{y = y_1\} \cap \hat{\gamma}$ and $\hat{x}_2 = \{y = y_2\} \cap \hat{\gamma}$. Notice that the whole interval with the end points (x_1, y_1) and $(\hat{x}_1, y_1)$ belongs to $\mathbf{S}_1$, as well as the interval with end points (x_2, y_2) and $(\hat{x}_2, y_2)$. Now we have

(34) $|\log a(x_1, y_1) - \log a(x_2, y_2)| \leq$
$$\leq |\log a(x_1, y_1) - \log a(\hat{x}_1, y_1)| + |\log a(\hat{x}_1, y_1) - \log a(\hat{x}_2, y_2)| + |\log a(\hat{x}_2, y_2) - \log a(x_2, y_2)|$$

Due to the Cone condition the width of $\mathbf{S}_1$ is not greater than $200C_0^4h^{1+\nu}$. By the mean value theorem we have

(35) $|\log a(x_1, y_1) - \log a(\hat{x}_1, y_1)| \leq$
$$\leq \frac{1}{|a(x_1^*, y_1)|}\left|\frac{\partial a}{\partial x}(x_1^*, y_1)\right||x_1 - \hat{x}_1| \leq C_1h^{1+\nu} \cdot C_1h^{-2-\nu} \cdot 200C_0^4h^{1+\nu} = 200C_1^2C_0^4h^\nu$$

Similarly

$$|\log a(\hat{x}_2, y_2) - \log a(x_2, y_2)| \leq 200C_1^2C_0^4h^\nu$$

Now parameterize the curve $\hat{\gamma}$ by the parameter y, $\hat{\gamma} = \hat{\gamma}(x(y), y), y \in [y_1, y_2]$ (or $y \in [y_2, y_1]$ if $y_2 < y_1$). Consider a function $g(y) = \log a(\hat{\gamma}(x(y), y)$. Since $\hat{\gamma}$ is tangent to the cone field $\{K^s\}$, for some $y^* \in [y_1, y_2]$ we have

(36) $|g(y_1) - g(y_2)| = |g'(y^*)| \cdot |y_1 - y_2| = \frac{1}{|a(\hat{\gamma}(x(y^*), y^*)|} \cdot \left|\frac{\partial a}{\partial x}\hat{\gamma}_x' + \frac{\partial a}{\partial y}\hat{\gamma}_y'\right| \cdot |y_1 - y_2| \leq$
$$\leq C_1h^{1+\nu}\left(\left|\frac{\partial a}{\partial x}\right||\hat{\gamma}_x'| + \left|\frac{\partial a}{\partial y}\right||\hat{\gamma}_y'|\right) \leq C_1h^{1+\nu}(C_1h^{-2-\nu} \cdot 100C_0^6h^{1+\nu} + C_1) \leq$$
$$\leq C_1h^\nu(100C_1C_0^6 + C_1h) \leq 200C_1^2C_0^6h^\nu$$

Finally we have

$$|\log a(x_1, y_1) - \log a(x_2, y_2)| \leq 400 C_1^2 C_0^4 h^\nu + 200 C_1^2 C_0^6 h^\nu < 600 C_1^2 C_0^6 h^\nu$$

$\square$

The following Proposition directly follows from Lemmas 10, 11, 15, 16, 17, 18, and 19.

Proposition 20. *The map* $\mathbf{T} : \mathbf{S}_0 \cup \mathbf{S}_1 \to \mathbf{S}$ *belongs to the class* $\mathcal{F}(C^*, \gamma, \varepsilon)$ *(see definition 6), where* $C^* = 120 C_1^4 C_0^6$, $\gamma = 1200 C_1^3 C_0^6 h^\nu$ *and* $\varepsilon = 120 C_1^3 C_0^6 h^{1+\nu}$.

5.6. **Final step.**

Proof of Theorem 3. Properties 1. and 2. of Theorem 3 clearly follow from the construction and the Cone condition. Let us combine now Proposition 20 with Duarte Distortion Theorem (Theorem 5) and Lemma 14. Let us assume that h is small enough so that $e^{D(C^*, \varepsilon, \gamma)} < 2$. Then we have

$$\frac{1}{4} h^{-1} \leq \tau_L(K^s) \leq 4 h^{-1}, \quad \frac{1}{200} C_0^{-4} h^\nu \leq \tau_R(K^s) \leq 500 C_0^4 h^\nu,$$

$$\frac{1}{4} h^{-1} \leq \tau_L(K^u) \leq 4 h^{-1}, \quad \frac{1}{200} C_0^{-4} h^\nu \leq \tau_R(K^u) \leq 500 C_0^4 h^\nu.$$

Therefore

$$\tau_L(K^s) \tau_R(K^s) \geq \frac{1}{800} C_0^{-4} h^{-1+\nu} \to \infty \quad \text{as} \quad h \to 0 \quad (\text{i.e. } a \to -1).$$

Similarly $\tau_L(K^u) \tau_R(K^u) \to \infty$ as $a \to -1$, so $\tau_{LR}(\Lambda) \to \infty$ as $a \to -1$. Notice that this implies that Lemma 1 can be applied to the Cantor sets K^s and K^u, and this is how Duarte proved existence of the conservative Newhouse phenomena in [Du2, Du3].

To check the property 4., we apply Proposition 4 (notice that we are exactly in the setting of Remark 2). We have

$$(37) \quad \dim_H K^s \geq \frac{\log\left(1 + \frac{\tau_L(K^s)}{1 + \tau_R(K^s)}\right)}{\log\left(1 + \frac{1 + \tau_L(K^s)}{\tau_R(K^s)}\right)} \geq \frac{\log\left(1 + \frac{\frac{1}{4} h^{-1}}{1 + 500 C_0^4 h^\nu}\right)}{\log\left(1 + \frac{1 + 4 h^{-1}}{\frac{1}{200} C_0^{-4} h^\nu}\right)} =$$

$$= \frac{\log\left[h^{-1}\left(h + \frac{1}{4 + 2000 C_0^4 h^\nu}\right)\right]}{\log\left[h^{-1-\nu}\left(h^{1+\nu} + 200 C_0^4 (4 + h)\right)\right]} = \frac{1 - O((\log h)^{-1})}{1 + \nu - O((\log h)^{-1})} > \frac{1}{1 + 2\nu}$$

if h is small enough. Since ν could be chosen arbitrary small, $\dim_H K^s \to 1$ as $h \to 0$ (i.e. $a \to -1$). Similarly $\dim_H K^u \to 1$ as $a \to -1$. Since $\dim_H \Lambda = \dim_H K^s + \dim_H K^u$ ([MM], see also [PV]), we have

$$\dim_H \Lambda \to 2 \quad \text{as} \quad a \to -1.$$

Theorem 3 is proved. $\square$

6. Conservative homoclinic bifurcations and hyperbolic sets of large Hausdorff dimension: the proof

Here we derive Theorem 4 from Theorem 3.

Proof of Theorem 4. First of all, Theorem B from [Du4] claims that a generic unforlding of a conservative homoclinic tangency leads to appearance of C^2-stably-wild hyperbolic basic set. More precisely, there exists an open set $\mathcal{U}_0 \subset \mathbb{R}^1, 0 \in \overline{\mathcal{U}_0}$, such that

- each map $f_\mu, \mu \in \mathcal{U}_0$, has a basic set Λ_μ^* exhibiting homoclinic tangencies;
- $P_\mu \in \Lambda_\mu^*$, where P_μ is a continuation of the saddle P_0;
- there exists a dense subset $D_0 \subset \mathcal{U}_0$ such that for each $\mu \in D_0$ the saddle P_μ has a quadratic homolcinic tangency which unfolds generically with μ.

Choose a sequence of parameter values $\{\mu_n\}_{n\in\mathbb{N}} \subset D_0$ dense in $\mathcal{U}_0$. Fix any $a \in \mathbb{R}$. The renormalization technics by Mora-Romero [MR] prove that an appropriately chosen and rescaled map near a homoclinic tangency is C^r–close to a Henon map H_a. Namely, the following statement holds.

Theorem 8 ([MR], based on [AS, GS2])**.** *Let $\{f_\nu\} \subset \mathrm{Diff}^\infty(M^2, \omega)$ be a smooth family of area preserving maps unfolding generically a quadratic homoclinic tangency at the point $Q_0 \in M$ and parameter $\nu = 0$. Then there are, for all large enough $n \in \mathbb{N}$, reparametrizations $\nu = \nu_n(a)$ of the parameter variable ν and a-dependent coordinates*

$$(x, y) \mapsto \Psi_{n,a}(x, y) \in M^2$$

such that

(1) for each compact K, in the (a, x, y)-space, the images of K under the maps

$$(a, x, y) \mapsto (\nu_n(a), \Psi_{n,a}(x, y))$$

converge to $(0, Q_0) \in \mathbb{R} \times M^2$, as $n \to \infty$.

(2) the domains of the maps

$$(a, x, y) \mapsto (a, \Psi_{n,a}^{-1} \circ f_{\nu_n(a)}^n \circ \Psi_{n,a}(x, y))$$

converge to $\mathbb{R}^3$ as $n \to \infty$ and the maps convege in the C^∞ topology to the conservative Henon map

$$(a, x, y) \mapsto (a, y, -x + a - y^2).$$

By Theorem 3 for a slightly larger than -1 the map H_a has an invariant hyperbolic set Λ_a of Hausdorff dimension close to 2 with persistent hyperbolic tangencies. By continuous dependence of Hausdorff dimension of an invariant hyperbolic set on a diffeomorphism [MM, PV] near each μ_n there is an open interval of parameters $U_n \subset \mathcal{U}_0$ such that for $\mu \in U_n$ the map f_μ has an invariant locally maximal transitive hyperbolic set Δ_μ^* with Hausdorff dimension greater than $2 - \delta$. Set $\mathcal{U} = \cup_{n\in\mathbb{N}}U_n$ and $D = D_0 \cap \mathcal{U}$. The hyperbolic saddle P_μ and the set Δ_μ^* are homoclinically related, see Lemma 2 from [Du2]. Therefore for every $\mu \in U_n$ there exists a basic set Δ_μ such that $P_\mu \in \Delta_\mu$ and $\Lambda_\mu^* \cup \Delta_\mu^* \subset \Delta_\mu$. Since Λ_μ^* has persistent homoclinic tangencies, so does Δ_μ. Also, $\dim_H \Delta_\mu \geq \dim_H \Delta_\mu^* > 2 - \delta$. By construction the parts (1) and (2) of Theorem 4 are now satisfied.

Let us now observe how elliptic periodic points appear. Take any $\mu \in \mathcal{U}$. If Q_μ is a transversal homoclinic point of the saddle P_μ then in can be continued for some intervals of parameters $I_Q \subseteq \mathcal{U}$. Assume that $I_Q \subseteq \mathcal{U}$ is a maximal subinterval of $\mathcal{U}$ where such a continuation is possible. All homoclinic points of P_μ for all values $\mu \in \mathcal{U}$ generate countable number of such subintervals $\{I_s\}_{s\in\mathbb{N}}$ in $\mathcal{U}$.

From [MR] it follows that for each I_s there exists a residual set $R_s \subseteq I_s$ of parameters such that for $\mu \in R_s$ the corresponding homoclinic point Q_μ is an accumulation point of elliptic periodic points of

f_μ. Denote $\widetilde{R}_s = \left(\mathcal{U}\backslash\overline{I_s}\right) \cup R_s$ – residual subset of $\mathcal{U}$. Now set $\mathcal{R}_1 = \cap_{s\in\mathbb{N}}\widetilde{R}_s$ – also a residual subset in $\mathcal{U}$. For $\mu \in \mathcal{R}_1$ every transversal homoclinic point of the saddle P_μ is an accumulation point of elliptic periodic points of f_μ, and this proves (3.1).

Now let us see that for a residual set of parameters in $\mathcal{U}$ the homoclinic class of P_μ has full Hausdorff dimension. In the same way as we constructed $\mathcal{U}$ starting with $\mathcal{U}_0$, from Theorem 3 and [MR] it follows that for every $m \in \mathbb{N}$ there exists an open and dense subset $\mathcal{A}_m \subset \mathcal{U}$ such that for every $\mu \in \mathcal{A}_m$ there exists a hyperbolic set Δ_μ^m such that $\dim_H\Delta_\mu^m > 2 - \frac{1}{m}$. From Lemma 2 from [Du2] it follows that P_μ and Δ_μ^m are homoclinically related. Therefore there exists a basic set $\widetilde{\Delta}_\mu^m$ such that $P_\mu \in \widetilde{\Delta}_\mu^m$ and $\Delta_\mu^m \subset \widetilde{\Delta}_\mu^m$. In particular, for $\mu \in \mathcal{R}_2 = \cap_{m\geq 1}\mathcal{A}_m$ we have $\dim_H H(P_\mu, f_\mu) = 2$. Set $\mathcal{R} = \mathcal{R}_1 \cap \mathcal{R}_2$. This proves (3.2).

The last property (3.3) follows from the following Lemma.

Lemma 21. *Let $\Lambda \subset M^2$ be a basic set of a surface diffeomorphism. Then*

$$\dim_H\{x \in \Lambda|\ \mathcal{O}^+(x) \text{ is dense in } \Lambda \text{ and } \mathcal{O}^-(x) \text{ is dense in } \Lambda\} = \dim_H\Lambda.$$

Indeed, take any $\mu \in \mathcal{R}$, and consider the set $S_\mu = \{x \in H(P_\mu, f_\mu) \mid P_\mu \in \omega(x) \cap \alpha(x)\}$. By construction of $\mathcal{R}$, for any $m \in \mathbb{N}$ the homoclinic class $H(P_\mu, f_\mu)$ contains a hyperbolic set $\widetilde{\Delta}_\mu^m$ such that $P_\mu \in \widetilde{\Delta}_\mu^m$ and $\dim_H\widetilde{\Delta}_\mu^m > 2 - \frac{1}{m}$. Due to Lemma 21, the set of points in $\widetilde{\Delta}_\mu^m$ whose positive and negative semiorbits are both dense in $\widetilde{\Delta}_\mu^m$ has Hausdorff dimension greater than $2 - \frac{1}{m}$, and, hence, $\dim_H S_\mu > 2 - \frac{1}{m}$. Since this is true for arbitrary large m, we have $\dim_H S_\mu = 2$.

Lemma 21 is a folklore, and can be seen as a corollary of results from [Ma, MM].

This completes the proof of Theorem 4.   □

7. ON HAUSDORFF DIMENSION OF STOCHASTIC SEA OF THE STANDARD MAP: THE PROOF

In this section we derive the results on the Hausdorff dimension of stochastic sea of the standard map (Theorems 1 and 2) from the result on hyperbolic sets of large Hausdorff dimension appearing after a conservative homoclinic bifurcation (Theorem 4).

In the study of the standard family in the current context Duarte [Du1] proved the following important results:

Theorem A (Duarte, [Du1]). *There is a family of basic sets Λ_k of f_k such that:*

1. Λ_k is dynamically increasing, meaning for small $\varepsilon > 0$, $\Lambda_{k+\varepsilon}$ contains the continuation of Λ_k at parameter $k + \varepsilon$.

2. Hausdorff Dimension of Λ_k increases up to 2. For large k,

$$\dim_H(\Lambda_k) \geq 2\frac{\log 2}{\log\left(2 + \frac{9}{k^{1/3}}\right)}.$$

3. Λ_k fills in $\mathbb{T}^2 \ni (x, y)$, meaning that as k goes to ∞ the maximum distance of any point in $\mathbb{T}^2$ to Λ_k tends to 0. For large k, the set Λ_k is δ_k-dense on $\mathbb{T}^2$ for $\delta_k = \frac{4}{k^{1/3}}$.

Theorem B (Duarte, [Du1]). *There exists $k_0 > 0$ and a residual set $R \subseteq [k_0, \infty)$ such that for $k \in R$ the closure of the f_k's elliptic points contains Λ_k.*

Theorem C (Duarte, [Du1]). *There exists $k_0 > 0$ such that given any $k \geq k_0$ and any periodic point $P \in \Lambda_k$, the set of parameters $k' \geq k$ at which the invariant manifolds $W^s(P(k'))$[6] and $W^u(P(k'))$ generically unfold a quadratic tangency is dense in $[k, +\infty)$.*

Theorem 1 should be considered as an improvement of Theorems A and B.

Proof of Theorem 1 and Theorem 2. We begin with the following technical statement. Denote by $\mathcal{N}(N) = (n_1, \ldots, n_N)$ an N-tuple with $n_i \in \mathbb{N}$.

Proposition 22. *There exists $k_0 > 0$ such that for each $N \in \mathbb{N}$ there is a family of finite open intervals $\mathcal{U}_{\mathcal{N}(N)} \subseteq [k_0, +\infty)$ indexed by N-tuples $\mathcal{N}(N) = (n_1, \ldots, n_N)$ satisfying the following properties:*

U1: *For pair of tuples $\mathcal{N}(N) \neq \mathcal{N}'(N)$ intervals $\mathcal{U}_{\mathcal{N}(N)}$ and $\mathcal{U}_{\mathcal{N}'(N)}$ are disjoint.*

U2: *For any tuple $\mathcal{N}(N+1) = (\mathcal{N}(N), n_{N+1})$ we have $\mathcal{U}_{\mathcal{N}(N+1)} \subseteq \mathcal{U}_{\mathcal{N}(N)}$.*

U3: *The union $\cup_{n_1 \in \mathbb{N}} \mathcal{U}_{n_1}$ is dense in $[k_0, +\infty)$, and for each $N \in \mathbb{N}$ the union $\cup_{j \in \mathbb{N}} \mathcal{U}_{(\mathcal{N}(N), j)}$ is dense in $\mathcal{U}_{\mathcal{N}(N)}$.*

U4: *Every diffeomorphism f_k, $k \in \mathcal{U}_{\mathcal{N}(N)}$, has a sequence of invariant basic sets*

$$\Lambda_k^{(n_1)} \subseteq \Lambda_k^{(n_1, n_2)} \subseteq \ldots \subseteq \Lambda_k^{\mathcal{N}(N)},$$

and $\Lambda_k^{\mathcal{N}(N)}$ depends continuously on $k \in \mathcal{U}_{\mathcal{N}(N)}$.

U5: *$\Lambda_k \subseteq \Lambda_k^{(n_1)}$ for each $n_1 \in \mathbb{N}$ and $k \in \mathcal{U}_{n_1}$, where Λ_k is a hyperbolic set from Theorem A.*

U6: *$\dim_H \Lambda_k^{\mathcal{N}(N)} > 2 - 1/N$.*

U7: *For any point $x \in \Lambda_k^{\mathcal{N}(N)}$ there exists an elliptic periodic point p_x of f_k such that $\mathrm{dist}(p_x, x) < 1/N$.*

Proof of Proposition 22. Notice that Theorem 4 and Theorem C directly imply the following statement:

Lemma 23. *Given $k^* \in (k_0, +\infty)$, $\varepsilon > 0$ and $\delta > 0$, there exists a finite open interval $V \subset (k^* - \varepsilon, k^*)$ such that for all $k \in V$ the map f_k has a basic set Λ_k^* such that*

1) Λ_k^* depends continuously on $k \in V$;

2) $\Lambda_k^* \supseteq \Lambda_k$, where Λ_k is a basic set from Theorem A;

3) Hausdorff dimension $\dim_H \Lambda_k^* > 2 - \delta$,

4) For any point $x \in \Lambda_k^*$ there exists an elliptic periodic point p_x of f_k such that $\mathrm{dist}(p_x, x) < \delta$.

Proposition 22 can be reduced to Lemma 23. Indeed, let us show how to construct the intervals $\mathcal{U}_{n_1}$ and the sets $\Lambda_k^{(n_1)}$. Let $\{k_l\}_{l \in \mathbb{N}}$ be a dense set of points in $(k_0, +\infty)$. Apply Lemma 23 to each $k^* = k_l$, $l \in \mathbb{N}$, for $\delta = 1$, $\varepsilon = \varepsilon_l < \frac{1}{l}$. That gives a sequence of open intervals $\{V_l\}_{l \in \mathbb{N}}$. Since the sequence $\{k_l\}_{l \in \mathbb{N}}$ is dense in $(k_0, +\infty)$ and $\varepsilon_l \to 0$, intervals $\{V_l\}$ are dense in $(k_0, +\infty)$.

Take $\mathcal{U}_1 = V_1$. If $\mathcal{U}_1, \ldots, \mathcal{U}_t$ are constructed, take V_s – the first interval in the sequence $\{V_l\}_{l \in \mathbb{N}}$ that is not contained in $\overline{\cup_{n_1=1}^{t} \mathcal{U}_{n_1}}$. Then $V_s \backslash \overline{\cup_{n_1=1}^{t} \mathcal{U}_{n_1}}$ is a finite union of K open intervals. Take

---

[6]Recall that $P(k')$ denotes the continuation of the periodic saddle P at parameter k'.

those intervals as $\mathcal{U}_{t+1}, \ldots, \mathcal{U}_{t+K}$, and continue in the same way. This gives a sequence of a disjoint intervals $\{\mathcal{U}_{n_1}\}_{n_1 \in \mathbb{N}}$ with desired properties.

Now, assume that intervals $\{\mathcal{U}_{\mathcal{N}(N)}\}$ are constructed. Take one of the intervals $\mathcal{U}_{\mathcal{N}(N)}$. Due to Theorem C, the set $\Lambda_k^{\mathcal{N}(N)}$ exhibits persistent tangencies. Therefore, application of Theorem 4 gives a dense sequence of intervals $\{V_{\mathcal{N}(N),l}\}_{l \in \mathbb{N}}$ in $\mathcal{U}_{\mathcal{N}(N)}$ such that for each $k \in V_{\mathcal{N}(N),l}$ the map f_k has a basic set Δ_k such that Hausdorff dimension $\dim_H \Delta_k > 2 - \dfrac{1}{N+1}$ and $\Delta_k \cap \Lambda_k^{\mathcal{N}(N)} \neq \emptyset$.

Now we need the following lemma from hyperbolic dynamics.

Lemma 24. *Let Δ_1 and Δ_2 be two basic sets (i.e. locally maximal transitive hyperbolic sets) of a diffeomorphism $f : M^2 \to M^2$ of a surface M^2 that are homeomorphic to a Cantor set. Suppose that $\Delta_1 \cap \Delta_2 \neq \emptyset$. Then there is a basic set $\Delta_3 \subseteq M^2$ such that $\Delta_1 \cup \Delta_2 \subseteq \Delta_3$.*

Remark 8. *Having in mind some possible generalizations, we notice that Lemma 24 holds also for higher dimensional diffeomorphisms (two-dimensionality of the phase space is not used in the proof).*

Proof of Lemma 24. Due to recent result of Anosov [An] any zero-dimensional hyperbolic set is contained in a locally maximal hyperbolic set. Therefore in our case $\Delta_1 \cup \Delta_2$ is contained in some locally maximal hyperbolic set $\widetilde{\Delta}$. Spectral decomposition theorem claims that $\widetilde{\Delta}$ is a finite disjoint union of basic sets. One of these basic sets must contain Δ_1, and since $\Delta_1 \cap \Delta_2 \neq \emptyset$, the same basic set has to contain Δ_2. $\qquad\square$

Apply Lemma 24 to Δ_k and $\Lambda_k^{\mathcal{N}(N)}$, and denote by $\widetilde{\Lambda}_k^{\mathcal{N}(N)} \supset \Delta_k \cup \Lambda_k^{\mathcal{N}(N)}$ the corresponding basic set. The set $\widetilde{\Lambda}_k^{\mathcal{N}(N)}$ also has persistent tangencies. The unfolding of a homoclinic tangency creates elliptic periodic orbits which shadow the orbit of homoclinic tangencies. The creation of these generic elliptic points can be seen from the renormalization at conservative homoclinic tangencies, see [MR]. Shrinking $V_{(\mathcal{N}(N),l)}$ if necessary we can guarantee that $\widetilde{\Lambda}_k^{\mathcal{N}(N)}$ can be $\frac{1}{N+1}$-accumulated by elliptic periodic points. Now the same procedure that we applied above to intervals $\{V_l\}$ gives a collection of disjoint intervals $\{\mathcal{U}_{(\mathcal{N}(N),n_{N+1})}\}_{n_{N+1} \in \mathbb{N}}$ in $\mathcal{U}_{\mathcal{N}(N)}$. For any $k \in \mathcal{U}_{(\mathcal{N}(N),n_{N+1})} \subset V_{(\mathcal{N}(N),l)}$ we take $\Lambda_k^{(\mathcal{N}(N),n_{N+1})} = \widetilde{\Lambda}_k^{\mathcal{N}(N)}$. By construction, all the properties in Proposition 22 are now satisfied. $\qquad\square$

Now let us explain how Theorems 1 follows from Proposition 22. Set $\mathbf{U}_N = \cup_{\mathcal{N}(N)} \mathcal{U}_{\mathcal{N}(N)}$. Due to **U3)** the set $\mathbf{U}_N$ is dense in $[k_0, +\infty)$. Therefore $\mathcal{R} = \cap_{N \in \mathbb{N}} \mathbf{U}_N$ is a residual subset of $[k_0, +\infty)$. Properties **U1)** and **U2)** imply that for each $k \in \mathcal{R}$ the value k belongs to each element of the uniquely defined nested sequence of intervals

$$\mathcal{U}_{n_1} \supseteq \mathcal{U}_{n_1,n_2} \supseteq \ldots \supseteq \mathcal{U}_{\mathcal{N}(N)} \supseteq \ldots$$

Therefore for $k \in \mathcal{R}$ the sequence of basic sets

$$\Lambda_k \subseteq \Lambda_k^{(n_1)} \subseteq \Lambda_k^{(n_1,n_2)} \subseteq \ldots \subseteq \Lambda_k^{\mathcal{N}(N)} \subseteq \ldots$$

is defined such that Hausdorff dimension $\dim_H \Lambda_k^{\mathcal{N}(N)} > 2 - 1/N$. Since k is fixed now, redenote $\Lambda_k^N = \Lambda_k^{\mathcal{N}(N)}$. Items 1.– 3. of Theorem 1 follows from **U5)** and **U6)**.

The closure of the union of a nested sequence of transitive sets is transitive, so property 4. follows.

For a locally maximal transitive invariant hyperbolic set of a surface diffeomorphism the Hausdorff dimension of the set is equal to the Hausdorff dimension of any open subset of this set, see [MM]. This implies the property 5. for the sets $\Omega_k, k \in \mathcal{R}$.

Property 6. follows directly from **U7)**.

Finally, in order to prove Theorem 2 it is enough to consider the family of basic sets $\Lambda_k^{\mathcal{N}(N)}$ defined for $k \in \mathbf{U}_N$ for large enough N. $\qquad\square$

References

[ABC]   Arnaud M.-C., Bonatti C., Crovisier S., Dynamiques symplectiques génériques, *Ergodic Theory Dynam. Systems,* vol.25 (2005), no. 5, pp. 1401–1436.

[A]   Alexeyev V., Sur lallure finale du mouvement dans le problme des trois corps. (French) Actes du Congrs International des Mathmaticiens (Nice, 1970), Tome 2, pp. 893–907, Gauthier-Villars, Paris, 1971.

[An]   Anosov D.V., Extension of 0-dimensional hyperbolic sets to locally maximal ones, to appear in *Matematicheskij Sbornik*.

[AS]   Afraimovich V., Shilnikov L., On critical sets of Morse-Smale systems, *Trans. Moscow Math. Soc.*, v. 28 (1973), pp. 179-212.

[BL]   Bloor K., Luzzatto S., Some remarks on the geometry of the Standard Map, *International Journal of Bifurcation and Chaos*, vol. 19 (2009), pp. 2213–2232.

[BC]   Bonatti C., Crovisier S., Récurrence et généricité. *Invent. Math.,* vol.158 (2004), no. 1, pp. 33–104.

[Bu]   Bunimovich L., Mushrooms and other billiards with divided phase space, *Chaos*, vol. 11 (2001), pp. 802–808.

[BG]   Brännström N., Gelfreich V., Asymptotic Series for the Splitting of Separatrices near a Hamiltonian Bifurcation, arXiv:0806.2403v1 [math.DS], 2008.

[Ch]   Chernov V., On separatrix splitting of some quadratic area preserving maps of the plane, *Regular and Chaotic Dynamics* vol. 3, no.1 (1998), pp. 49–65.

[C]   Chirikov B., A universal instability of many-dimensional oscillator systems, *Phys. Rep.* 52 (1979), pp. 263 – 379.

[DH]   Dankowicz H., Holmes P., The existence of transverse homoclinic points in the Sitnikov problem, *J. Differential Equations*, vol. 116 (1995), no. 2, pp. 468–483.

[Do]   Donnay V., Geodesic flow on the two-sphere I. Positive measure entropy, *Ergod. Theory Dynam. Syst.* vol. 8 (1988), pp. 531–553.

[DN]   Downarowicz T., Newhouse S., Symbolic extensions and smooth dynamical systems. *Invent. Math.,* vol. 160 (2005), no. 3, pp. 453–499.

[Du1]   Duarte P., Plenty of elliptic islands for the standard family of area preserving maps, *Ann. Inst. H. Poincare Anal. Non Lineaire,* vol. 11 (1994), no. 4, pp. 359–409.

[Du2]   Duarte P., Abundance of elliptic isles at conservative bifurcations, *Dynamics and Stability of Systems*, vol. 14, no. 4, (1999), pp. 339–356.

[Du3]   Duarte P., Persistent homoclinic tangencies for conservative maps near the identity, *Ergod.Th & Dynam. Systems*, (2000), **20**, pp. 393–438.

[Du4]   Duarte P., Elliptic Isles in Families of Area Preserving Maps, *Ergod.Th & Dynam. Systems*, vol. 28 (2008), pp. 1781–1813.

[FS1]   Fontich E., Simo C., The splitting of separatrices for analitic diffeomorphisms, *Ergodic Theory and Dynamical Systems* vol.10 (1990), pp. 295–318.

[FS2]   Fontich E., Simo C., Invariant manifolds for near identity differentiable maps and splitting of separatrices, *Ergodic Theory and Dynamical Systems* vol.10 (1990), pp. 319–346.

[GP]   Garcia A., Perez-Chavela E., Heteroclinic phenomena in the Sitnikov problem, *Hamiltonian systems and celestial mechanics* (Patzcuaro, 1998), pp. 174–185, *World Sci. Monogr. Ser. Math., 6*, World Sci. Publ., River Edge, NJ, 2000.

[GL]   Gelfreich V., Lazutkin V., Splitting of separetrices: perturbation theory and exponential smallness, *Russ. Math. Surveys* **56** 3, 499–558;

[G1]   Gelfreich V., Splitting of a small separatrix loop near the saddle-center bifurcation in area-preserving maps, *Physica D* 136 (2000), no. 3-4, pp. 266–279.

[G2]   Gelfreich V., A proof of the exponentially small transversality of the separatrices for the standard map, *CMP* 201 (1999), pp. 155–216.

[G3]   Gelfreich V., Conjugation to a shift and the splitting of invariant manifolds, *Applicationes Mathematicae* 24 (1996), no. 2, pp. 127–140.

[GSa]   Gelfreich V., Sauzin D., Borel summation amd splitting of separatrices for the Henon map, *Ann.Inst.Fourier, Grenoble* 51, 2 (2001), pp. 513–567.

[GS1]   Gonchenko S., Shilnikov L., On two-dimensional area-preserving maps with homoclinic tangencies that have infinitely many generic elliptic periodic points, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* 300 (2003), *Teor. Predst. Din. Sist. Spets.*, Vyp. 8, pp. 155–166, 288–289; translation in *J. Math. Sci. (N. Y.)* 128 (2005), no. 2, pp. 2767–2773.

[GS2]   Gonchenko S., Shilnikov L., Invariants of Ω-conjugacy of diffeomorphisms with a non-generic homoclinic tragectory, *Ukr. Math. J.*, v.28 (1990), pp. 134 – 140.

[GST]   Gonchenko S., Turaev D., Shilnikov L., Homoclinic tangencies of arbitrarily high orders in conservative and dissipative two-dimensional maps, *Nonlinearity* vol. 20 (2007), no. 2, pp. 241–275.

[GK1]   Gorodetski A., Kaloshin V., Conservative homoclinic bifurcations and some applications, *Steklov Institute Proceedings*, vol. 267, pp. 76–90.

[GK2]   Gorodetski A., Kaloshin V., How often surface diffeomorphisms have infinitely many sinks and hyperbolicity of periodic points near a homoclinic tangency, *Advances in Mathematics*, vol. 208 (2007), pp. 710–797.

[GHK]   Gorodetski A., Hunt B., Kaloshin V., Newton interpolation polynomials, discretization method, and certain prevalent properties in dynamical systems, Vol. 2, *Proceedings of ICM 2006, Madrid, Spain,* European Math Society (2006) pp. 27–55.

[Go]   Goroff D., Hyperbolic sets for twist maps, *Ergodic Theory Dynam. Systems*, vol. 5 (1985), no. 3, pp. 337–339.

[H]   Henon M., Numerical study of quadratic area preserving mappings, *Quarterly of Applied Mathematics*, (1969), vol. 27, no.3, pp. 291–312.

[IL]   Ilyashenko Yu., Li Weigu, Nonlocal bifurcations, *Providence, R.I., American Mathematical Society,* 1999.

[I]   Izraelev F., Nearly linear mappings and their applications, *Physica D*, vol. 1 (1980), no. 3, pp. 243–266.

[L]   Rafael de la Llave, A tutorial on KAM theory, In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 175-292, Amer. Math. Soc., Providence, RI, 2001.

[Li]   Liverani C., Birth of an elliptic island in a chaotic sea, *Math. Phys. Electron. J.* 10 (2004).

[LS]   Llibre J., Simo C., Oscillatory solutions in the planar restricted three-body problem, *Math. Ann.* **248**, (1980), pp. 153–184.

[Ma]   Manning A., A relation between Lyapunov exponents, Hausdorff dimension and entropy, *Ergodic Theory Dynam. Systems*, vol. 1 (1982), no. 4, 451–459 .

[MM]   Manning A., McCluskey H., Hausdorff dimension for horseshoes, *Ergodic Theory Dynam. Systems* , vol. 3 (1983), no. 2, pp. 251–260.

[MMP]   MacKay R. S., Meiss J. D., Percival I. C., Stochasticity and transport in Hamiltonian systems, *Phys. Rev. Lett.*, vol. 52 (1984), no. 9, pp. 697–700.

[MR]   Mora L., Romero R., Moser's invariant curves and homoclinic bifurcations, *Dynamic Systems and Applications*, (1997), vol. 6, pp. 29–42.

[Mo]   Moreira G., Stable intersections of Cantor sets and homoclinic bifurcations, *Ann.Inst.Henri Poincare* no.13 (6), (1996), pp. 741–781.

[Mos]   Moser J., Stable and random motions in dynamical systems, Princeton University Press, 1973.

[McG]   McGehee R., A stable manifold theorem for degenerate fixed points with applications to celestial mechanics, *J. Diff. Equations*, vol. 14 (1973), 70–88.

[N1]   Newhouse S., Non-density of Axiom A(a) on S^2, *Proc. A.M.S. Symp. Pure Math.*, **14**, (1970), pp. 191–202.

[N2]   Newhouse S., Diffeomorphisms with infinitely many sinks. *Topology* **13**, (1974), pp. 9–18.

[N3]   Newhouse S., Lectures on dynamical systems, *Dynamical systems (C.I.M.E. Summer School, Bressanone, 1978), Progr. Math., 8, Birkhuser, Boston, Mass.*, 1980, pp. 1–114.

[N4]   Newhouse S., The abundance of wild hyperbolic sets and nonsmooth stable sets for diffeomorphisms, *Inst. Hautes tudes Sci. Publ. Math.*, No. 50 (1979), pp. 101–151.

[N5]   Newhouse S., Topological entropy and Hausdorff dimension for area preserving diffeomorphisms of surfaces, *Dynamical systems, Vol. III—Warsaw, Asterisque,* no. 51, *Soc. Math. France, Paris,* (1978), pp. 323–334.

[NP]   Newhouse S., Palis J., Cycles and bifurcation theory, *Asterisque*, vol. 31 (1976), pp. 44–140.

[PT1]   Palis J., Takens F., Hyperbolicity and sensitive chaotic dynamics at homoclinic bifurcations. Cambridge University Press, 1993.

[PT2]   Palis J., Takens F., Hyperbolicity and the creation of homoclinic orbits, *Ann. of Math.* (2) 125 (1987), no. 2, pp. 337–374.

[PV]   Palis J., Viana M., On the continuity of Hausdorff dimension and limit capacity for horseshoes. *Lecture Notes in Math., 1331, Springer, Berlin,* 1988.

[PY]   Palis J., Yoccoz J.-Ch., Non-Uniformly Hyperbolic Horseshoes Arising from Bifurcations of Poincare Heteroclinic Cycles, arXiv:math/0604174v1 [math.DS].

[P]   Pesin Ya., Characteristic Lyapunov exponents and smooth ergodic theory, *Uspehi Mat. Nauk*, vol. 32 (1977), no. 4 (196), pp. 55–112.

[Pr]   Przytycki F., Examples of conservative diffeomorphisms of the two-dimensional torus with coexistence of elliptic and stochastic behavior, *Ergod. Theory Dynam. Syst.*, vol. 2 (1982), pp. 439–463.

[R]   Robinson C., Bifurcations to infinitely many sinks, *Comm. Math. Phys.*, **90**, (1986), no. 3, pp. 433–459.

[Sin]   Sinai Ya., Topics in ergodic theory, *Princeton Mathematical Series*, 44 (1994). Princeton University Press, Princeton, NJ, viii+218 pp.

[Si]   Sitnikov K., The existence of oscillatory motions in the three-body problems. *Dokl. Akad. Nauk SSSR* 133, pp. 303–306 (Russian); translated as *Soviet Physics. Dokl.* 5 (1960), pp. 647–650.

[SS]   Shepelyansky D., Stone A., Chaotic Landau level mixing in classical and quantum wells, *Phys. Rev. Lett.* vol. 74 (1995), pp. 2098–2101.

[SM]   Siegel C., Moser J., Lectures on Celestial Mechanics, *Springer: Berlin-Heidelberg-New York*, 1971.

[S]   Sternberg S., The structure of local homeomorphisms of euclidian n-space, III *Amer. J. Math.* vol.81 (1959), pp. 578–604.

[TY]   Tedeschini-Lalli L., Yorke J. How often do simple dynamical processes have infinitely many coexisting sinks? *Comm. Math. Phys.* **106**, (1986), no. 4, pp. 635–657.

[W]   Wojtkowski M., A model problem with the coexistence of stochastic and integrable behavior, *Commun. Math. Phys.*, vol. 80 (1981), pp. 453–464.

[X1]   Xia J., Melnikov method and transversal homoclinic points in the restricted three-body problem, *J. of Diff. Equations*, vol 96, 1 (1992), pp. 170–184.

[X2]   Xia J., Some of the problems that Saari did not solve, *Celestial Mechanics, Dedicated to Donald Saari for his 60th Birthday. Proceedings of an International Conference on Celestial Mechanics, 15-19 December, 1999 at Northwestern University, Evanston, Illinois*, Providence, RI: American Mathematical Society, Contemporary Mathematics, vol. 292 (2002), p. 267.

Department of Mathematics, University of California, Irvine CA 92697, USA

E-mail address: asgor@math.uci.edu